





RESULT	3
LOCUS	E13599
DEFINITION	DNA encoding a signal chain Fv antibody of anti-gp130 antibody gpX7.
ACCESSION	E13599
NID	93352401
KEYWORDS	JP 1997220092-A/2.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 720)
AUTHORS	Eki-da,T., Yasukawa,K., Imanaka,T. and Takagi,M.
TITLE	PRODUCTION OF SINGLE-STRAND EV ANTIBODY
JOURNAL	Patent: JP 1997220092-A 2 26-Aug-1997;
COMMENT	TOSOH CORP OS None OC Artificial sequences. PN JP 1997220092-A/2 PD 26-Aug-1997 PF 15-FEB-1996 JP 1996027622 PI EKI DA TEIJI,I, YASUKAWA KIYOSHI, IMANAKA TADAYUKI, PI TAKAGI MASAHIRO PC C12N1/09,C12N1/21,C12P21/08,(C12N1/21,C12R1:19),(C12P21/08, PC C12R1:19); CC strandedness: Double; CC topology: Linear; CC hypothetical: No; CC anti-sense: No; FH key Location/Qualifiers FT source . 1..720 /organism='Artificial sequences'. FEATURES source location/Qualifiers 1..720 /organism='unidentified' /db_xref='taxon:32644'
BASE COUNT	169 a 193 c 190 g 168 t
ORIGIN	
Query Match	64.3%; Score 458; DB 6; Length 720;
Best Local Similarity	80.2%; Pred: 8.8e-131;
Matches 576; Conservative 1; Mismatches 131; Indels 10; Gaps 3;	
Oy	4 GACCAGGACTATGCCAGCTTCACACAATCCGTGTCATCTCCAGGGAGAAGGTACAC 63
Db	1 GAGCTCGTGTCAACCAAGCTCCAGCATTTGAATATCTGCATCTCCAGGGAGAAGGT-CAC 59
Oy	64 AATGACTTGCAAGGGCACCCCAAGTAGTAAGTATACATGCATCGATGATWACAGAACCCAGG 123
Db	60 CATGACCTGCAATATGCAAGCTCAAGTCAAGTTACTTCATGATTTGTGTACACAGAAACCCAG 119
Oy	124 ATCTCTCCCCCAAACCTTGGATTATACCAACATCCAACTCKGTTGTGAGTCCCTGCTCG 183
Db	120 ATCTCTCCCCCAAACCTTGGATTATCTCACTCAACCAACGCGCTCTGAGTCCCTCTCG 179
Oy	184 CTTCAGTGGCGGTGGGTCTGGGAACTCTTACTCTCTCACAGTCAGACAGAGTGGAGCTGA 243
Db	180 CTCACGTGGCAATGGGTCTGGGAGCAACTTACTCTTCACAAATCAGACAGCTTGGAGCTGA 239
Oy	244 AGATCTGCACTTATTTACTGTCAGACAGTGAAGTGGTAGCCACCCAGCTTGGAGGGGG 303
Db	240 AGATCTGCACTTATTTACTGTCAGACAGTGAAGTGAATAACCCGCTACACTTGGGTGCTGG 299
Oy	304 GTCCAAAGCTGAAAATAAAGTTTCACTCTGTGTTCTGGTAATCTTCTGAAGGTAAA-- 362
Db	300 GACCAAGCTCGAGCTGGGTGGGTGGCTCGGAGCGGTGGTGGGTGGGTGGGTGGGAGATC 359
Oy	362 -GGTGTGACGCTGCAGAGATTCAGAGCTTAAGTGGTGAAGCCTGGAGGTTCATGAAGAT 420

[illegible]

QY	352	TGAAGTAAAGTGTCACCTGAGGAGTCAAGACTGAGGTGGTAGAAGCCTGGAGTTTC	411
Db	360	CAAAAGCACGCACAGTCAAACCTGAGSAGCTGATCTGACTGGTGAAGCCTGGAGCTTC	419
QY	412	AATGAAGATATCTCGCAAGACTTCTGGTTACTCATATTCACTGGCCACACATGAATGGGT	471
Db	420	AGTGAAGCTCTCTGCAGAAGCTTTTGGCTACATTTACACAGTACTAGGATGACCTGGGT	479
QY	472	GAAACAGACCCATGGAAGAACCCTTGAGTGGATTGGACTTTATTCCTTCAATGGTGA	531
Db	480	GAACACAAGGCAATGAGCAAGGCCCTTGAGGATTTGGAATATTTATCTCGTGGTGGTAG	539
QY	532	TACTAACTAACCAACAAAGTTCAAGGGCAAGGCCACATTACTGTAGACAAGTGTCCAG	591
Db	540	TACTAACTAGATGAGAAGATTCAAGACAAAGCCACACTGACTGTAGACATCTCCAG	599
QY	592	CACAGCCTACATSGAGACTCCTCAGTCTGACATCTGAGGACTCTSCAGTCTATTACTGTGC	651
Db	600	CACAGCCTACATGCACCTCAGCAGCCTGACATCTGAGGACTCTGCGGCTATTACTGTAC	659
QY	652	AAGGAGGGTTACGACTGTGACTCTGCATG-----TCTGGGGCCGCAAGGACACGGT	702
Db	660	AAGGAGGAGCCGGAACTGGGTCTACTATGCTATGACTACTGACTACTGGGGTCAAGGAACCTCAGT	719
QY	703	CACCGTCTCC	712
Db	720	CACCGTCTCC	729
RESULT	5		
EI0361	ID	EI0361	standard; RNA; ROD; 732 BP.
XX	AC	EI0361:	
XX	NI	d1108698	
DT	08-OCT-1997	(Rel. 52, Created)	
DT	08-OCT-1997	(Rel. 52, Last updated, Version 1)	
XX	cDNA encoding a monoclonal antibody against human Interleukin 2 receptor gamma chain.		
DE	XX	JP 1995313188-A/1.	
OS	Mus sp.		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia;		
OC	Sciurognathi; Muridae; Murinae; Mus.		
RN	[1]		
RP	1-732		
RA	Shimamura T., Hamuro J., Nakazawa H., Kanayama Y., Sugamurra K.,		
RA	Takeshita T.;		
RT	"IMMUNOSUPPRESSIVE AGENT";		
RL	Patent number JP 1995313188-A/1, 05-DEC-1995.		
RL	AJINOMOTO CO INC, SUGAMURA KAZUO.,		
XX			
OS	Mus sp. (mouse)		
CC	PN	JP 1995313188-A/1	
CC	PD	05-DEC-1995	
CC	PR	21-APR-1994 JP 1994082836	
CC	PF	21-APR-1993 JP 93P 94491, 07-MAR-1994 JP 94P 36065	
CC	PI	SHIMAMURA TOSHIAKI, HAMURO JUNJI, NAKAZAWA HARUMI,	
CC	PI	KANAYAMA YUKA,	
CC	PI	SUGAMURA KAZUO, TAKESHITA TOSHITCHI	
CC	PC	C12P21/08,A61K39/395,A61K39/395,C12N1/21,C12N5/20,	
CC	PC	C12N15/13//C12N15/06,	
CC	PC	(C12P21/08,C12R1:19),(C12P21/08,C12R1:91),(C12N1/21,C12R1:19);	
CC	CC	strandedness: Double;	
CC	CC	topology: Linear;	
CC	*Key	Location/Qualifiers	
CC	FH		

CC	FT	source	1..732	/organism="Mus sp."
CC	FT		/cell_type="hybridoma"	
CC	FT		/cell_line="Gp-2"	
CC	FT	mat_peptide	1..732	/product="anti-IL-2 receptor gamma chain"
CC	FT			
XX	XX	Key	Location/Qualifiers	
FH	FT	source	1..732	
FT	FT		/organism="Mus sp."	
XX	XX			
Sequence	732 BP; 180 A; 204 C; 182 G; 166 T; 0 other;			
Query Match	61.0%: Score 434; DB 27; Length 732;			
Best Local Similarity	78.6%: Pred. No. 2,2e-123;			
Matches	574; Conservative 1; Mismatches 136; Indels 19; Gaps			
QY	1	ATGGAACCTGAGCTGACCCAGTCTCCAGACATCCCTGTGCAATCTCCAGGGGAGAGTA	60	
Db	1	ATGGACATCTGCTGACACCAAGTCTCCATCAATCAATGATGTGATCTCTAGGGGA-ACGGGT	59	
QY	61	CACAATGACTTGAGAGGCCACCCCAAGTGA-AGTATGATGACTGTATACGA	114	
Db	60	CACCACTGACCTGACATGCTGACAGTCAAGTGAATGCTCAGTGTCTGATACAGCA	119	
QY	115	GAGGCAAGATCCCTCCCAAACTTGATGATTTATACACATCAACCTGCTTGAGAT	174	
Db	120	GAGGCAAGATCTCTCCCAAACTTGATGATTTATGACATCAACCTGCTTGAGAT	179	
QY	175	CCCTGCTGCTTCAAGTGGCGGTGGTCTGGGACCTTACTCTCAGAGTACAGAGAT	234	
Db	180	CCCACTGCTTCAAGTGGCGGTGGTCTGGGACCTTACTCTCAGATACAGAGAT	239	
QY	235	GAGGCTGAAGATGCTGGCATTATTAATGACAGAGTGGAGTCTGACCCCAACGTT	294	
Db	240	GAGGCTGAAGATGCTGGCATTATTAATGACAGAGTGGAGTCTGACCCCAACGTT	299	
QY	295	CGAGGAGGGGCTCCAAAGCTGGAATAAAGGTCTCA-CCCTGGTCTGTAATCTTC	351	
Db	300	CGGTGCTGGGACCAAGCTGGAATAAAGGTCTCA-CCCTGGTCTGTAATCTTC	359	
QY	332	TGAAGCTAAAGTGTGACAGCTGACAGAGTCAAGACCTGAGTGTGTAAGCTTGAGTTC	411	
Db	360	CAAAAGCGACGAGTCAAACTCGAGAGTCTGGAATCTGAGCTGGAGGCTGGAGCTTC	419	
QY	412	AATGAAGATTCCTGCAAGCTCTGGTATCACTTCACTGCGCCACACCAATCAACGGGT	471	
Db	420	AGTGAAGCTGTCTGCAAGGCTTCTGGCTACACATTCACACAGCTCTGATGATCAGTGGT	479	
QY	472	GAGCAGAGCCATGAGAAACCTTGAAGTGTGATGATGATTAATCTTCAATGCTGA	531	
Db	480	GAGCAGAGCCATGAGAAACCTTGAAGTGTGATGATGATTAATCTTCAATGCTGA	539	
QY	532	TACTACTACCAACCGAAGTTCAGAGGGAAGGCCACATTTATGTAAGCAATGCTGCG	591	
Db	540	TACTACTACCAAGTTCAGAGGGAAGGCCACACACTGATGTGAGACATCTCTCCAG	599	
QY	592	CACAGCTACATGAGACTCTCAGTCACTGACATCACTAGAGACTGTGAGTCTATTAAGTGTG	651	
Db	600	CACAGCTACATGAGACTCTCAGAGAGCTGACATCACTGAGGACTGTGGGTCTATTAAGTGTG	659	
QY	652	AAAGAGGCTTACGAGCTGTGATCTGATG-CTTGAGGCGCAGGAGCAACGGT	702	
Db	660	AAAGAGGCGCCGGAAGTGGTCTACTATGATGACTACTGAGGCTCAAGGAACCTCACT	719	
QY	703	CACGCTGCTCC 712		
Db	720	CACGCTGCTCC 729		



132407	LOCUS	132407	720 bp	DNA	PAT	07-JAN-1997
DEFINITION	Sequence 2 from patent US 5587458.					
ACCESSION	132407					
NID	g1823198					
KEYWORDS	.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 720)					
TITLE	King,C.Richter, Kasprzyk,P.G. and Bird,R.E.					
JOURNAL	Anti-erbB-2 antibodies, combinations thereof, and therapeutic and					
FEATURES	diagnostic uses thereof					
SOURCE	Patent: US 5587458-A 2 24-DEC-1996;					
	location/Qualifiers					
	1..720					
BASE COUNT	184 a 178 c 179 g 179 t					
ORIGIN	/'organism="unknown"					

Query Match	59.3%	Score 422;	DB 6;	length 720;
Best Local Similarity	79.3%	Pred. No. 1, 1e-119;		
Matches 511; Conservative	1;	Mismatches 131;	Indels 1;	Gaps 1;

OY	8	TGCAGCTGACCCAGCTCTCCACCAATCCGTCGTGATCTCCAGGGGAAAGAGTACACATG	67
Db	2	TGCAGCTGACCCAGCTCTCCACCAATCAATGTCGTGATCTCCAGGGGAAAGAGT-CACATG	60
OY	68	ACTTCGAGGGCCACCCCAAGTGTGAATTACATGCATCGTATCGCAGAAAGCCAGATCC	127
Db	61	ACCTTCAGTGGCAGCTCAAGGTGAAGTAAACATGCACTGGATATCGCAGAAAGTCAAGCACC	120
OY	128	TCGCCCAAACTTGGATTTATACCAACATCCAACTKCGTCTGGAGGCCGCGCGCTTC	187
Db	121	TCGCCCAAACTCTGGGTTTATGAACATCCAACTGGGTTCTGGAGTCCCAAGGTCGCTTC	180
OY	188	AGTGGCGGTGGGTCTGGGACCTCTTACTCTCTCCAGCTCAGCAGAGTGGAGGCTGAAGAT	247
Db	181	AGTGGCAGTGGGTCTGGAAACTCTTACTCTCTCCAGATCAGCAGACATGAGAGGCTGAAGAT	240
OY	248	GCTGCACATTATTCATGCCAGCAGTGGAGTCTGAGCCCAACCACGTTGGAGGGGGCTCC	307
Db	241	GCTGCACATTATTTATTTGTTATCAGAGGGAGTGGATCCATTCAGCTTCGCGTGGGGACA	300
OY	308	AAGCTGGAAATAAAAAGGTTTACGCTGAGTCTGGTAACTTCTGAAAGGTAAGGCTG	367
Db	301	AAGTTGGAAATAAAAAGGTTTCACTCCGAGTCTGGTAATTTCTGAAAGGTAAGGCTG	360
OY	368	CAGCTCAGAGGATCAGSACCTGAGGTTGGTGAAGCCTGGAGGTTCAATGAAGATATCTGC	427
Db	361	CAGCTCAGAGGATCTGGGGTTGAGGTTGTCGAGAGAGGGGCTTAAGTCAAGTTGCTGC	420
OY	428	AAGACTTTGGTTACTCAATTCACTGCGCCACACATGAACCTGGTGAAGACAGCATGA	487
Db	421	AAAGTTTGTGACTTCACATTAAAGACATATATATCCACTGGGTGAAGCAGAGGCTCGAA	480
OY	488	AAGAACCTTGATGGATTTGGAAGTATTAATCTTAACATGGTGTACTTAATCACTAACACAG	547
Db	481	CAGGGCTTGGATGGATTTGATGGATTTATCTCTGGAATGGTAATACTGTAATGACCCG	540
OY	548	AAGTTCAGAGGCAAGGCCACATTTACTGTAGACAAGTGTCCACAGACGCTTACATGAG	607
Db	541	AAATTCAGAGGCAAGGCCAGATATAACAGACAGACATCCTCAACGGGGGCTACTTCAG	600
OY	608	CTTCTCAGTCTGACATCTGAGAGCTCTGCAGTCTAATTAATCTGCG	651
Db	601	CTCAGCAGCTCTGACATCTGAGAGCACTGCGCTTAATTAATCTGCG	644

RESULT	7			
I45604				
LOCUS	i45604	916 bp	DNA	PAT
				29-JUL-1997

DEFINITION	Sequence 29 from patent US 5637481.
ACCESSION	145604
NID	92469706
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 916) Iedebetter,J.A., Gilliland,L.K., Hayden,M.S., Linsley,P.S., Bajorath,J. and Bell,H.Perry.
TITLE	Expression vectors encoding bispecific fusion proteins and methods of producing biologically active bispecific fusion proteins in a mammalian cell
JOURNAL	Patent: US 5637481-A 29 10-JUN-1997;
FEATURES	Location/Qualifiers
SOURCE	1..916
BASE COUNT	254 a 223 c 228 g 211 t
ORIGIN	/organism="unknown"

Query Match	50.0%	Score 356.2;	DB 6,	Length 916;
Best Local Similarity	71.9%	Pred. No. 2.3e-99;		
Matches 557; Conservative	0;	Mismatches 173;	Indels 37;	Gaps 4

OY	2	TGGACCTCCACCTACACCAGCTCCACAGCAATCCGTGGCATCTCCAGGGGAGAGGTAC	61
Db	74	TCGACATCCAAATATACACAGACTACATCTCCCTGCTGGCTCTCTGGGAGACGAGT -C	1322
OY	62	ACAATGACTTTCAGAGGGCCACC---CAAGTGAATTACATGACACTGTTATCAGCAAG	118
Db	133	ACCATTCAGTTCCAGGGCAGTCAAGACATTCGGCATTTATTTAACTGGTATCAGCAAAA	1322
OY	119	CCAGGATCTCCCCCAAACTTGGATTATPACCATCCAACTCGCTTGTGAGTCCCT	178
Db	193	CCAGTGGGAATGTTAAACTCTGATCTACTACACATCAAGATTATACACTCAGGAGTCCA	2522
OY	179	GCTGCTTCATGGCGGCGGTGCTGGGACCTTACTCTCTCACAGTCACAGAGTGGAG	238
Db	253	TCGAGTTCACTGGCAGTGGGTCTGGAAACAGATTATTTCTTCACCATTTGGCAAC	312
OY	239	GCTGAAGATGCTGCACATTTATCTGGCAGAGAGTGGAGTGTAGCCACACAGTTGGA	298
Db	313	CCAGAGATATTTGCCACTTACTTTTGGCACAAGGGTATACGCTTCCGTGGACGTTGGT	372
OY	299	GGGGGGTCCAAGCTGGAATTTAAAGTTTCACTCTGTGTTCTGTGTAATCTTTCGAAGT	358
Db	373	GGAGGCACCAAACTGTTAACCAAAAGGGAGCTGGGTGGCGTGGCGGTGGGTGGG	422
OY	359	AAAGT-----GTGCAGCTGACGAGAGTCAGGACCTGAGAGTGGTG	397
Db	433	TCGGGTGGCGCGGATCATCGATGAGAGTCCAGCTGCACAAGTGTGGACCTGAACTGGT	492
OY	398	AAAGCTGGAGTTCAATGAGATATCTGTCAAGACTTGTGTTACTATTCATCTGGCCAC	457
Db	493	AAAGCTGGAGTTCAATGAGATATCTGTCAAGAGCTCTGTGTTACTATTCATCTGGCCAC	552
OY	458	ACCATGAATCGGGTGAAGCAGAGCATGGAAGAAACCTGTAGTGGATTTGGACTATTAAT	517
Db	553	ATTCGTAATCGGCTGTAAGCAGAGCATGGAAGAAACCTGTAGTGGATTTGGACTATTAAT	612
OY	518	CTTTCATGATGATAGTAACTACACACAGAAAGTTCAAGGGCAGAGCCACTTACTGTA	577
Db	613	CCATTCAAAAGTCTTACTACTACCAACAGAAATTCAGAGGGCAAGGCCACTTAACTGTA	677
OY	578	GACAAGTCTGCAGACAGCCTACATGAGAGCTCTCAAGTCTGACATCTGAGAGACTGCA	637
Db	673	GACAAGTCTACACACAGCCTACATGAGAGCTCTCAAGTCTGACATCTGAGAGACTGCA	732
OY	638	GTCATTACTGTGGAAG-----GAGGTTACGAGAGTGGTACTTCGATGTGTGG	685
Db	733	GTCATTACTGTGGAAGTCTGGTACTAGTGGTACTCTGGAGCTGGTACTTCGATGTGTGG	792





RESULT	11	SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
LOCUS	SYNAR3IGA					
DEFINITION		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
ACCESSION		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
KEYWORDS		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
SOURCE		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
ORGANISM		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
REFERENCE		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
AUTHORS		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
TITLE		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
JOURNAL		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
MEDLINE		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
FEATURES		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
source		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
CDS		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
misc-feature		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
BASE COUNT		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
ORIGIN		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
Query Match		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
Best Local Similarity		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
Matches		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
479; Conservative		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
1; Mismatches		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
225; Indels		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
4; Gaps		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990
2;		SYNAR3IGA	738 bp	DNA	SYN	15-SEP-1990

QY	418	GATATCCGCAAGACTTCTGGTTACTACTTCACTTCACTGCGCACACACATGAACCTGGGTGACGA	477
DB	423	ACTCTCCGTGCAGGCTCTGGATTACTTTCACTAGTGGCTTTGGAAAGCACTGGGTTCGCA	482
QY	478	GAGCATCTGGAAGAAGCACTTGAAGTGAATGGATTTATCTTATCCTTCAATGGGATACTAA	537
DB	483	GCGTCCAGAGAAGAGGGGCTGGAGTGGGTGGCATACTTTAGTGGCGGTAAATACATCTA	542
QY	538	CTACACACAGAAGTTCAGAGGCGAACAGCCACATTTACTGTAGACAAGTGTCCAGCAGAC	597
DB	543	CTATGCAACACAGTAGAAGGCGCGATTCACCATCTCCAGAGACAACTCCAAAGAACCCCT	602
QY	598	CTACATGAGAGCTCCCTCACTTGACATCTGAGAGACTCTGCAGCTTATTTACTGTGCAAGAG	657
DB	603	GTTCTGTGAAGATGACCACTCTAAAGGTCTGAGAGACAGGCGCAATGTATTTCTGTGCAAGATC	662
QY	658	GATTACGAGACTGTACTTCAGATGTGTGGGGCGGACAGGACCAAGCTGCACC	706
DB	663	CCATTACTACGGCTACTTCTATGATGTAGACTTACTGGGGTCAGGGAACC	711
RESULT	12		
LOCUS	AR012822	797 bp	DNA
DEFINITION	Sequence 3 from patent US 5763733.		PAT
ACCESSION	AR012822		
NID	93971140		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 797)		
AUTHORS	Whitlow,M., Filipula,D. and Shorr,R.		
TITLE	Antigen-binding fusion proteins		
JOURNAL	Patent: US 5763733-A 3 09-JUN-1998;		
FEATURES	Location/Qualifiers		
Source	1..797		
BASE COUNT	196 a	203 c	195 g
ORIGIN	196 a	203 c	195 g
	203 t		
Query Match	44.8%	Score 319;	DB 6;
Best Local Similarity	72.0%;	Pred. No. 6.7e-88;	Length 797;
Matches 447;	Conservative 1;	Mismatches 161;	Indels 12;
			Gaps 2;
QY	95	TACATGACACTGGTATACAGACAGAGCAAGATCCCTCCCAAAACCTTGGATTATACACA	154
DB	112	TACTTGGCTGTGATACACAGAAACAGGCGACGTCTCTTAACTCTGTTACTGGGCA	171
QY	155	TCCAACTKGCCTTCGTGAGAGTCCCTGCTGCTTCACTGAGCGGTGGTCTGGGACCTTTAC	214
DB	172	TCCGGTAGGAATCTGGGGTCCCTGATGCTTACACAGGACGATCGGAGACATTTTC	231
QY	215	TCTCTACAGTACAGAGAGTGGAGGCTTAAGATGCTGCCACTTATTACTCCAGCAGTGG	274
DB	232	ACTCTCTCATCAGCAGAGTGTGAAGACTGAAGACCTGCGCAGTTATTACTGTACGACATAT	291
QY	275	AGTGTACCCACACCCAGCTGTGGAGGGGGTCCAAAGCTGGAATAAAGGTTCTACCTCT	334
DB	292	TATATACTTATCCCTTACGTTGTGGTCTGGGACCAAGCTTGTGCTAAAGAGCTTATCTTCC	351
QY	335	GATTTCGTAAATCTTCTGGAAGTAAAGT--GTGACAGCTGACAGAGTACAGAGACTGAG	391
DB	352	GGTACGGGCAAACTCTCTGGAAGCAAAAGTCAAGTTCACGTACAGCTGACAGCTGAG	411
QY	392	GTGGTGAAGCCTGAGGTTCAATGAAGATATCTGCAAGACTTCTGTTACTATTCACT	451
DB	412	TTGGTGAACCTCGGGGCTTCACTAGTGAAGATTTCTTCAAGGCTTCTGCTACACCTTCACT	471
QY	452	GCCACACACATGAATCTGGGTGAAGCAGACCATGGAAGAAGACCTTGAATGATTTGACTT	511
DB	472	GACCAATGAATTCATCTGGGTGAAGAAGAACCTTGAAACAGGGCGCTGGAATGATTTGGATAT	531

0y	512	ATTATTCCTTACAAATGGTGATACTACTACTACACCAAGAACTTAAAGGCAAGGCCACATT	571
0y	512	ATTATTCCTTACAAATGGTGATACTACTACTACACCAAGAACTTAAAGGCAAGGCCACATT	571
Db	532	TTTTTCCCGGAAATGAATGATTTTAAATACATGAGAGGTTAAAGGCCACACTG	591
0y	572	ACTGTAGACAAAGTCGTCCGACAGACCTTACATGAGAGCTCTCAGTCTGACATCTGAGGAC	631
Db	592	ACTGCAGACAAATCTCTCCAGCAGCTGCTTACGTGACAGCTTACAGCCTGACATCTGAGAT	651
0y	632	TCGTGAGCTTATTAAGTGTGCAAGAGAGGCTTACGAGCTGTGACTTGCATGTCTGAGGCGCA	691
Db	652	TCGTGAGCTTATTAAGTGTGCAAGAGAGGCTTACGAGCTGTGACTTGCATGTCTGAGGCGCA	702
0y	692	GGGACACGGTCAACGCTTCC	712
Db	703	GGACCTCGGTCAACGCTTCC	723
RESULT 13			
LOCUS	AR012824	803 bp	DNA
DEFINITION	Sequence 7 from patent US 5763733.	PAT	04-DEC-1998
ACCESSION	AR012824		
NID	93971142		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 803)		
AUTHORS	Whitlow, M., Filpula, D. and Shorr, R.		
TITLE	Antigen-binding fusion proteins		
JOURNAL	Patent: US 5763733-A 7 09-JUN-1998;		
FEATURES	Location/Qualifiers		
Source	1..803		
	/organism="unknown"		
BASE COUNT	198 a	205 c	196 g
ORIGIN		204 t	
Query Match 44.6%; Score 317.8; DB 6; Length 803;			
Best Local Similarity 73.9%; Pred. No. 1.6e-87;			
Matches 416; Conservative 1; Mismatches 143; Indels 3; Gaps			
0y	95	TACATGCACTGGTATACAGAGAACGAGATCCTCCCAACCTTGATTTATACACA	154
Db	181	TACTTGCGCTGTGACAGACAGAAACAGGCGAGTCTCTTAACCTGTGATTTACTGGGCA	240
0y	155	TCCAACTKGTCTTGAGAGTCCCTGCTCCTTCAAGTGGCGGTGCTGGACCTTAC	214
Db	241	TCCGATAGGAAATCGGGGTCCCTGATGCTTACAGAGGAGTGGATCTGGACAGATTTC	300
0y	215	TCTCTCACAGTACAGAGAGTGAAGGTGAAGATGCTGCCACTTATTACTGCAGACAGTGG	274
Db	301	ACTCTCTCATATGAGTGTGAAGACTGGCAAGCTGGCAGTTTATTAATCTGACAGATAT	360
0y	275	AGTCGAGCCACCCAGCTTGGAGGGGGGCTCAACCTGAAATAAAGTCTTACTACT	334
Db	361	TATAGCTATCCCTTACGTTGGTGTGTGGGACCAAGCTTGTGTGAAGGCTTACTTCC	420
0y	335	GTTTGTGTAATCTTCTGAAGGTAAAGT--GTGACAGCTGCAGAGTCAAGACCTGAG	391
Db	421	GGTAGCGGCAAAATCCTCTGAAGCAAGGTCAAGTTCACTGACAGCACTGACGCTGAG	480
0y	392	GTGGTGAACCTCGAGAGTCAATGAAGATATCCTGCAAGAACTTGTGTTACTCATCTACT	451
Db	481	TGTGTAAACCTGGGGCTTCAATGAAGATTTCTGTGAAGGTTCTGGTACACCTTCACT	540
0y	452	GGCACACCATGAACTGGGTGGAAGACCACTGAAAGAAACCTTGATGATTTGACTT	511
Db	541	GACCAATGCAATTCATCGGTGGAACAGAACCTGTAACAGGGCTTGAATATGATGATAT	600
0y	512	ATTATTCCTTACAAATGGTGATACTACTACTACACCAAGAACTTAAAGGCAAGGCCACATT	571

QY	601	TTTTCTCCCGGAATGATGATTTTAAATACATATGAGAGTTCAGGCGACACTG	660
QY	572	ACTGTAGCAAGTCTGCTCCAGCACAGCCCTACATGAGACTCTCATGCTGACATCTGAGAC	631
Db	661	ACTGACAGCAATCTCTCCAGCAGCTGCTACAGGAGCTGACAGCCTGACATCTAGGAT	720
QY	632	TCTGCACTCTATCTGTCGACG	654
Db	721	TCTGCACTCTATCTGTCGACG	743
RESULT	14		
LOCUS	E13598	735 bp	DNA
DEFINITION	E13598	DNA encoding a single chain Fv antibody of anti-T3 antibody Ttl.	
ACCESSION	E13598		
KEYWORDS	JP 1997220092-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 735)		
AUTHORS	Ekida,T., Yasukawa,K., Imanaka,T. and Takagi,M.		
TITLE	PRODUCTION OF SINGLE-STRAND FV ANTIBODY		
JOURNAL	Patent: JP 1997220092-A 1 26-AUG-1997;		
COMMENT	TOSOH CORP		
OS	None		
OC	Artificial sequences.		
PN	JP 1997220092-A/1		
PD	26-AUG-1997		
PF	15-FEB-1996 JP 1996027622		
PI	Ekida,Tsuyoshi, Yasukawa Kiyoshi, Imanaka Tadayuki, Pi Masahiro		
PC	C12N1/09, C12N1/21, C12P21/08, (C12N1/21, C12R1:19), (C12P21/08,		
CC	C12R1:19);		
CC	screenedness: Double;		
CC	topology: Linear;		
CC	hypothetical: No;		
CC	anti-sense: No;		
FH	Key	Location/Qualifiers	
FT	source	1..735	
FT	Location/Qualifiers	1..735	
FEATURES	source	Location/Qualifiers	
BASE COUNT	165 a 192 c 197 g 191 t		
ORIGIN			
Query Match	44.4%	Score 315.8	DB 6; Length 735;
Best Local Similarity	68.6%	Pred. No. 8.5e-87;	
Matches	499; Conservative	1; Mismatches 208; Indels 19; Gaps	4.
QY	4	GACCTGCAAGCTACCCAGCTCCACCAATCCCTGTCGATCTCCAGGGGAGAGATAC	63
Db	7	GACATTCAGCTACCCAGCTCCACCAATCCCTGTCGATCTCCAGGGGAGAGATAC	65
QY	64	AATGACTTCAGAGGCGCCCAAGTGTAAAG-----TACATGACCTGGTATCAGCAGAA	117
Db	66	TATCAGCTTCAGAGGCGCCCAAGTGTAAAG-----TACATGACCTGGTATCAGCAGAA	125
QY	118	GCCAGAGTCTCCCGCAAACTTGATTTATACCAATCCCAACCTTCGCTTGGAGTCCC	177
Db	126	GCCAGAGTCTCCCGCAAACTTGATTTATAGCAATCCCAACCTTCGCTTGGAGTCCC	185
QY	178	TGCTGCTTCAGTGGGAGGCTGGGAGCTTACCTCTACATCTCCAGTCCAGAGATGA	237
Db	186	AATGCTTCATTCAGGAGGCTGGGAGCTTACCTCTACATCTCCAGTCCAGAGATGA	245
QY	238	GAGTGAAGTGTGCACTTATCTGCGACAGTGGAGTGTGAGCCACCAACGTTGCG	297



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 09:34:42 ; Search time 341.74 Seconds

(without alignments)  
3178.154 Million cell updates/sec

Title: US-08-704-178-1

Perfect score: 1 ATGACCTGACGCTGACCA.....GGACACGCTCACCCTCC 712

Sequence: 1 ATGACCTGACGCTGACCA.....GGACACGCTCACCCTCC 712

Scoring table: IDENTITY\_NUC

Searched: 2002476 seqs, 762712212 residues

Database :

EST:  
1: em\_est1.\*  
2: em\_est2.\*  
3: em\_est3.\*  
4: em\_est4.\*  
5: em\_est5.\*  
6: em\_est6.\*  
7: em\_est7.\*  
8: em\_est8.\*  
9: em\_est9.\*  
10: gb\_est1.\*  
11: gb\_est2.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: gb\_est6.\*  
16: gb\_est7.\*  
17: gb\_est8.\*  
18: gb\_est9.\*  
19: gb\_est10.\*  
20: gb\_est11.\*  
21: gb\_est12.\*  
22: gb\_est13.\*  
23: gb\_est14.\*  
24: gb\_est15.\*  
25: gb\_est16.\*  
26: gb\_est17.\*  
27: gb\_est18.\*  
28: gb\_est19.\*  
29: gb\_est20.\*  
30: gb\_est21.\*  
31: gb\_est22.\*  
32: em\_est10.\*  
33: em\_est11.\*  
34: em\_est12.\*  
35: em\_est13.\*  
36: em\_est14.\*  
37: em\_est15.\*  
38: em\_est16.\*  
39: em\_est17.\*  
40: em\_est18.\*  
41: em\_est19.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	238.2	33.5	396	25	AA691311 vs14f01.r
2	190.4	26.7	344	22	AA592800 vs25g11.r

3	173.6	24.4	310	30	AA180569	AA180569 uc70f09.r
4	162.4	22.8	294	28	AA1007196	AA1007196 ua73c04.r
5	162	22.8	461	31	AT201426	AT201426 q573c06.x
6	146	20.5	387	22	AA569186	AA569186 mm30d10.s
7	128.4	18.0	345	25	AA710291	AA710291 vl53a04.r
8	119	16.7	431	25	AA710970	AA710970 vl93h09.r
9	118.6	16.7	393	11	T27593	T27593 EST100653 H
10	118	16.6	379	21	AA291381	AA291381 zt44g02.r
11	116.2	16.3	335	17	AA318377	AA318377 EST20620
12	114.2	16.0	413	17	AA301347	AA301347 EST14279
13	112.4	15.8	336	29	AA987559	AA987559 or83g01.s
14	110.8	15.6	403	18	AA345486	AA345486 EST51505
15	110.8	15.6	398	24	AA423447	AA423447 ve80s03.r
16	110.2	15.5	488	11	H25625	H25625 y148g05.r1
17	110.2	15.5	400	20	AA456778	AA456778 zw27f11.r
18	109.4	15.4	253	17	AA295941	AA295941 EST101165
19	109.4	15.4	333	17	AA300732	AA300732 EST13847
20	107.6	15.1	370	17	AA295093	AA295093 EST100400
21	106.4	14.9	363	18	AA367405	AA367405 EST78511
22	106.2	14.9	470	12	H62115	H62115 yu40h01.r1
23	106	14.9	442	24	AA405415	AA405415 zu56d02.r
24	105.8	14.9	328	17	AA295311	AA295311 EST100471
25	105.4	14.8	279	17	AA300571	AA300571 EST13661
26	105.4	14.8	345	18	AA335086	AA335086 EST39457
27	105.2	14.8	833	16	AA170256	AA170256 ms87g10.r
28	103.6	14.6	209	11	T28938	T28938 EST61186 Ho
29	102.8	14.4	349	18	AA379044	AA379044 EST91999
30	102.8	14.4	418	21	AA515239	AA515239 ng69c07.s
31	102.2	14.4	471	31	AT233978	AT233978 EST230666
32	101.8	14.3	372	25	AA581192	AA581192 nd38b11.r
33	101.4	14.2	382	17	AA295786	AA295786 EST100987
34	101	14.2	463	11	R69532	R69532 yj82d09.r1
35	100.8	14.2	287	21	AA496610	AA496610 zv38g06.r
36	100.4	14.1	210	11	R69482	R69482 yj83c03.r1
37	100.4	14.1	352	17	AA301261	AA301261 EST14181
38	100	14.0	420	30	AA120005	AA120005 uc25e04.r
39	99.8	14.0	432	31	AT268604	AT268604 q047a10.x
40	99.2	13.9	363	20	AA464313	AA464313 zx78c12.r
41	99	13.9	242	21	AA479857	AA479857 zu35d05.r
42	96.8	13.6	339	25	AA710249	AA710249 vt49e01.r
43	96.2	13.5	503	10	R28232	R28232 yh57e08.r1
44	96	13.5	291	18	AA377074	AA377074 EST89603
45	95.4	13.4	287	11	T27609	T27609 EST101034 H

## ALIGNMENTS

RESULT 1  
AA691311  
LOCUS  
DEFINITION vs14f01.r1 Barstead mouse irradiated colon MPR47 Mus musculus cdna  
clone 1138201 5' similar to gb:x67211 M.musculus rearranged  
Immunoglobulin kappa light chain (MOUSE); mRNA sequence.  
ACCESSION AA691311  
NID g2692247  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
1 (bases 1 to 396)  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Stepien,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine







RESULT	5
AI201426/c	
LOCUS	AI201426
DEFINITION	AI201426 461 bp mRNA EST 30-OCT-1998
SIMILAR	qsr3c06.x1 NCI_CGAP_Pt28 Homo sapiens cDNA clone IMAGE:1943722 3
COMMENT	similar to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN); contains Alu repetitive element; mRNA sequence.
ACCESSION	AI201426
NID	q3754032
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS	Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 461)
JOURNAL	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)

Query Match	22.8%	Score 162	DB 31	Length 461
Best Local Similarity	72.4%	Pred. No. 1.1e-40		
Matches 210; Conservative	0	Mismatches 80	Indels 0	Gaps 0

(C) 2006 Pearson Education, Inc.

RESULT	6
AA569186	
LOCUS	
DEFINITION	AA569186 387 bp mRNA
	mm30d10.s1 NC1:CGAP_L1p2 Homo sapiens cDNA clone IMAGE:1061583
	similar to gb:L02325 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN);
	mRNA sequence.
ACCESSION	AA569186
NID	G2342240
KEYWORDS	EST.
SOURCE	human.

Contact: Robert Strausberg, Ph.D.  
Tel.: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Kitzman, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/MLML at: [www-bio.linnl.gov/dbirp/image/image.html](http://www-bio.linnl.gov/dbirp/image/image.html)

**/organism="Homo sapiens"**  
**/note="Vector: pAMP10; mRNA made from liposarcoma, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research**

Query Match	20.5%	Score 146	DB 22	Length 387
Best Local Similarity	69.0%	Pred. No. 1.1e-35		
Matches 200	Conservative 0	Mismatches 90	Indels 0	Gaps 0

QY	365	GTGACGTCGCAGAGATCGAGACCTGAGGTGGTGAAGCCTGAGGTCCTAATGAATATATCC	42
Db	98	GTCCACCTGCGTGCACATCTGGGGGCTGAGGTGAAGAAAGCCTGGGGCCTCAGTGAACCTTTCC	15
QY	425	TGCAGACTTCTGGTTACTCATTCATCTGGCCACACACATGAAGTGGGTGAAGCAGACCAT	48
Db	158	TGTAAAGCATCTGGATTCACCCCTCAACACATGCATATGCACTGGGTGCGACAGGCCCTT	21
QY	485	GGAAAGAACCTTGATGGATTTGAGACTTATATCCTTATCATGGTGATACATACTAACAC	54
Db	218	GGCGAAGGGCTTGAATGGGTGGGAATGATTCATTTCTAGTATGGTTATATATAGTAAGCA	27
QY	545	CAGAAATTCAAAGGGCAAGGCCACATTTACTGTAGACAGTGTCCAGACAGCCTCATAG	60
Db	278	CAGAAATTCCAGGGCAGAGTCAACCATGACAGGAGACAGTCCAGAGACACAGTCTATATG	33
QY	605	GAGCCTCCTACATCTGACATCTGAGAGATCTCCACATCTTACTCTGTGCAAG	65
Db	338	GAGCTTACACCTGTAGATCTGAGGACACGGCCGTTATTTACTGTGGCAG	38

RESULT	7
AA710291	
LOCUS	
DEFINITION	
AA710291	345 bp mRNA EST 24-DEC-1997
v33a04.i1 Barstread mouse irradiated colon MRLB7 Mus musculus CDNA	
Clone 1166766 5' similar to gb:X02464.mal1 IG KAPPA CHAIN PRECURSOR	
V-LIV REGION (HUMAN);gb:j00560 mouse Ig kappa mRNA from mope21 &	
other myeloma mRNA 3' (MOUSE);, mRNA sequence.	
ACCESSION	

ACCESSION	AA710291
NID	g2720209

KEYWORDS	EST.
SOURCE	house mouse.

ORGANISM

## REFERENCE AUTHORS

100

TITLE  
JOURNAL

COMMENT

1  
2  
3  
4  
5  
6  
7

## FEATURES

### SOURCE

			/clone_11b="Bairstead mouse irradiated colon MFLRB7" /dev_stage="8 weeks" /lab_host="DH10B"	
BASE COUNT	95 a	81 c	88 g	81 t
ORIGIN				
Query Match	18.0%	Score 128.4;	DB 25;	Length 345;
Best Local Similarity	64.4%;	Pred. No. 3.8e-30;		
Matches 206;	Conservative 1;	Mismatches 112;	Indels 1;	Gaps 1.
QY	4	GACCTGCAGCTGACCCAGATCTCCAGCAATCCTGTCGTGATCCGAGGGGAGAGAGTTACAC	63	
Db	25	GACATTGTGGATGAGACCCAGATCTCCAAAATTCATGTCCATCATCAGCTGGAGACAGGGTTCAGC	84	
QY	64	AATGACTTTCGAGGGGCCACCCCAAGT-GTAAGTTACATGACTCTGGTATATGACGAGAAAGCCAG	122	
Db	85	GTCACCTTCGAAGCCACGATAGAAATGTGGGTACTCAATGATGCTGGTATCAAGAGAACCCAG	144	
QY	123	GATCCTCCGCCCAAACTTTGATTATTAACACATCCCAACCTTCGCTTGGAGTCCCTGTCTC	182	

QY	183	GCTCAGTGGGCGTGGGCTGGGACCTCTACTCTCTCAGATGACGAGAGTGGAGGCTG	24
Db	205	GCTTCACAGGACGATCGGATCTGGGACAGATTTCCTCTCTCACCATGACAAATGGACGCTG	26
QY	243	AAGATGCTGCCACTTATTACTGCGCAGAGTGGAGTCTGAGCCACCCACGTTGGAGGGG	302
Db	265	AAGACTTGGCGAGATATTCTCTGTCAGCAATATACACGCTTCCTACACGCTCGAGAGGG	324
QY	303	GGTCCAAAGCTGAAATATAAA	322
Db	325	GGACCAAGCTGGAATATAAA	344

## RESULT

AA710970  
LOCUS

### DEFINITION

ACCESSION  
NID  
\*\*\*\*\*

**KEYWORDS**  
**SOURCE**  
**ORGANISM**

ORGANIC

## REFERENCE AUTHORS

1

TITLE  
JOURNAL  
COMMENT

## COMMENT

## FEA TYPES

FEATURES  
SOURCE

```

/organism="Mus musculus"
/strain="C57BL/6J"
/Note="Organ: mammary gland; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site: 1: Not I;
Site: 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15'
TGTTCACATCTGAGGTGGAGCGCCCGAATGCTTTTCTTTTCTTTTCTTTT
T 3'1; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo.
/db_xref="taxon:10090"
/clone="1178753"
/clone_lib="Soares mouse mammary gland NBMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"

```

```

BASE COUNT      112 a      103 c      112 g      104 t
ORIGIN
Query Match      16.7%; Score 119; DB 25; Length 431;
Best Local Similarity 62.0%; Pred. No. 3.6e-27;
Matches 188; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```

```

QY 366 TCCAGCTGAGAGTACGACCTGAGTGTGAGCTGGAGTTCAATGAAATTCCT 425
DB 112 TCCAGTGTGACGTGAGCTGAGCTGAGAGAGCTGAGAGATCAATCTCT 171
QY 426 GCAAGACTTCTGTTACTGATTCACCTGACACACCATGAGTGGTGAAGACCATG 485
DB 172 GCAAGCTTCTGGTGTACTTACATCAACCTAGATGAGCTGGTGAAGACCATG 231
QY 486 GAAAGAACCTTAGTGATGATGATTAATTCCTTCAATGAGTGTACTAATCAAC 545
DB 232 GAAAGGCTTAAAGTGTGATGATGATTAACACCTACTGAGTGCACAAATATGCTG 291
QY 546 AGAAGTTCAGAGGACGACATTTACTGATGACAAAGTGCACACAGCTCATATG 605
DB 292 ATGACTTCAAGGAGCGTTCCTCTCTTTGAAACCTGCGACACATGCTATTTGC 351
QY 606 ACCTCTCAGTGTGACATCTGAGAGCTGACATGATTAATGAGAGGATTCAGG 665
DB 352 AGATCAACAACCTCAAAATGAGGACACGCTACATATTTCTGCCCCAATATCTTCA 411
QY 666 ACT 668
DB 412 ACT 414

```

```

RESULT 9
LOCUS T27593 393 bp mRNA EST 06-SEP-1995
DEFINITION ESI100653 Homo sapiens cDNA 5' end similar to immunoglobulin kappa
light chain, V region (GB:L01279) (HT:3043).
ACCESSION T27593
KEYWORDS NID 9609691
SOURCE EST.
ORGANISM human primer-M13 Reverse library=Human Pancreas.

```

```

REFERENCE
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulder,R.A.,
Bult,C.J., Lee,N., Kirsch,E.F., Weissstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W.,
Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.M.,

```

```

Glodek,A., Guelm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S.,
Kelley,J.M., Kilmek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M.,
Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T.,
Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L.,
Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Uterback,T.R.,
Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A.,
Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Fertile,A.,
Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M.,
Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H.,
Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.-F., Wing,J., Xu,C.,
Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A.,
Haseitine,W.A., Fields,C., Fraser,C.M., and Venter,J.C.
Initial Assessment of Human Gene Diversity and Expression Patterns
Based Upon 52 Million Basepairs of cDNA Sequence
Unpublished (1995)
Other ESTs: TNC24356
Contact: Venter, J.C.
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: tdbinfo@tdb.tigr.org
For clone availability, additional sequence and expression
information related to this EST, please contact the TIGR Database
(tdbinfo@tdb.tigr.org).
Location/Qualifiers
1..393

```

```

FEATURES
source

```

```

BASE COUNT      88 a      113 c      96 g      94 t      2 others
ORIGIN

```

```

Query Match      16.7%; Score 118.6; DB 11; Length 393;
Best Local Similarity 65.7%; Pred. No. 4.7e-27;
Matches 216; Conservative 1; Mismatches 107; Indels 5; Gaps 3;

```

```

QY 4 GACCTGAGCTGACCCAGCTTCACCAATCCGTGTCATCTCCAGGGAGAGTTAC 63
DB 53 GACATCCAGATGACCCAGCTTCACCTCCCGTGTGATCTGTAGAGACAGAT -CAC 111
QY 64 AATGACTTGCAGGGCCACCCCAAGTGTATTAATGCT---ACTGTCATCAGACAGCC 120
DB 112 CCTCACTTGGCCGGCAAGTCAAGCATTAAGCACCCTTTTAAATGATGATCAACAACCC 171
QY 121 AGGATCCCTCCCAACCTGATTTATACCAATCCAAACCGCTGTCGAGTCCCGC 180
DB 172 AGGACAGCCCTTAAGCTCTGATCTGTGTCATCAATTTGGAGAGTGGGTCCTCATC 231
QY 181 TCGCTTCAAGTGGCGGTGGTGTGGAGCTCTACTCTCAGACAGTCAAGAGTGGAGCC 240
DB 232 AAGGTCAGTGGCTGTGATCTGGAGACAGATTTCACTCAACATCAGACAGTCTCAACC 291
QY 241 TGAAGATGCTGCCACTTATTTACTGCCAGACATG -GATGTGTAAGCCACACGTTGGAG 299
DB 292 TGAAGATGTTGCACCTTACTACTCAACAGAGTTCACCTTATGTCGAGGTTGGCC 351
QY 300 GGGGTCCAAGTGAATTAAGGTTCT 328
DB 352 AAGGACCAAGGTGAGGTCAACGAGCT 380

```

```

RESULT 10
LOCUS AA291381 379 bp mRNA EST 08-AUG-1997
DEFINITION Z144902.r1 Soares ovary tumor NBHOT Homo sapiens cDNA clone 725234
5' similar to gp:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION
(HUMAN); mRNA sequence.
ACCESSION AA291381
KEYWORDS NID 91939359
SOURCE EST.
ORGANISM human.

```



Db 122 ATCAAGGTTGACCTGAGTGGACAGATTCACTGTCACCATCAGCTGCA 181  
QY 238 GCGTGAAGATGCTGCGCATTAATTAATGCGACAGTGAAGTCCACCCAGCTGG 297  
Db 182 ACCTGAATATTTGGAACTTACTACTGTCACAGAGTTCACAGTCCCTTGAGCTTGG 241  
QY 298 AGGGGGTCCAAAGCTGGAATTAAGTTTACTCTG 335  
Db 242 CCAAGGGNCCAAGTGGCAATCAACGACTGTGGCTG 279

RESULT 12  
AA301347 413 bp mRNA EST 18-APR-1997  
LOCUS ES114279 Testis tumor Homo sapiens cDNA 5' end similar to similar  
DEFINITION to immunoglobulin kappa light chain, V region (GB:101279), mRNA  
sequence.  
ACCESSION AA301347  
MID 91953680  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;  
Homo.  
REFERENCE 1 (bases 1 to 413)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,  
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Mali,C.,  
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,  
Glodok,A., Guelm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,  
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shiley,R.,  
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,  
Beonakik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,  
He,M.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,  
Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,  
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,  
Fraser,C.M., and Venter,J.C.  
JOURNAL Initial assessment of human gene diversity and expression patterns  
MEDLINE based upon 83 million nucleotides of cDNA sequence  
COMMENT Nature 377 (6547 Suppl), 3-174 (1995)  
96026280

TITLE  
JOURNAL  
MEDLINE  
COMMENT

CONTACT: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: akkerlavet@igrr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. 413  
/organism="Homo sapiens"  
/note="Organ: testis; Vector: pBluescript SK-; Site\_1:  
ECORI: Site\_2: XhoI"  
/db\_xref="ATCC (lnhst):191658"  
/db\_xref="taxon:9606"  
/clone\_id="Testis tumor"  
/sex="male"  
/dev\_stage="adult"  
<1. >413  
100 a 110 c 105 g 98 t

MRNA  
BASE COUNT  
ORIGIN

Query Match 16.0%; Score 114.2; DB 17; Length 413;  
Best Local Similarity 65.0%; Pred. No. 1.2e-25;  
Matches 214; Conservative 1; Mismatches 109; Indels 5; Gaps 3;  
QY 4 GACCTGAGCTGACCCAGTCTCAGCAATCTGTCTCATCTCCAGGGAGAGTACAC 63  
Db 78 GACATTCAGATGACCCAGTCTCCATCTCCCTGCTCATCTGAGGAGACAGAGT -CAC 136  
QY 64 AATGACTTCAGAGGCCACCCCA---AGTGAAGTAAATGACATGATATGACGAGAAC 120  
Db 137 CATCATCTGCGGGCTGGGACGACACATTAACCATTAATTAATGATATCAACGAAAC 196  
QY 121 AGATCTCTCCCAACCTGATTTATACACATCCACACTGCTGAGAGTCCCTGC 180  
Db 197 AGGAAAGCCCTTAAGTCTCTGATGATGATGATGATGATGATGATGATGATGATG 256  
QY 181 TCGCTTCACTGCGCGGTGGGTGGGACCTCTTACTCTCTCAGACTCAGAGAGTGAAGC 240  
Db 257 AAGTTCAGTGGCGGTGGATCTGGGACAGATTTCATCTCACCATCACCAGTCTGCAAC 316  
QY 241 TGAAGATGCTGCCATTAATTAATGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 299  
Db 317 TGAAGATTTTGAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 376  
QY 300 GGGGTCGAAGCTGGAATTAAGTCT 328  
Db 377 AAGGACCAAGTGGAAATCAAAAGACT 405

RESULT 13  
AA987559 336 bp mRNA EST 23-JUL-1998  
LOCUS or83901.s1 NCI-CGAP\_Lus Homo sapiens cDNA clone IMAGE:1602480 3'  
DEFINITION similar to gb:u87789 IG GAMMA-1 CHAIN C REGION (HUMAN);, mRNA  
sequence.  
ACCESSION AA987559  
MID 93172923  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 336)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT

CONTACT: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/HLN at:  
[www.dio.llnl.gov/bdrip/image/image.html](http://www.dio.llnl.gov/bdrip/image/image.html)

Insert Length: 801 Std Error: 0.00  
Seq primer: -40m13 fwd. RT from Amersham  
High quality sequence stop: 322.  
Location/Qualifiers  
1. 336  
/organism="Homo sapiens"  
/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; 1st strand cDNA was prepared from a  
neuroendocrine lung carcinoma, and was then primed with a  
Not I - Oligo(dT) primer. Double-stranded cDNA was ligated  
to Eco RI adaptors (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of the modified pTR3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo.

/db\_xref="taxon:9606"  
/clone="IMAGE:1602480"  
/clone\_lib="NCI CGAP Lu5"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"

BASE COUNT 82 a 86 c 100 g 68 t  
ORIGIN

Query Match 15.8%; Score 112.4; DB 29; Length 336;  
Best Local Similarity 61.7%; Pred. No. 4.1e-25;  
Matches 179; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

OY 365 GTGACGCTGACAGAGTACAGGACCTGAGGTGAGAGCCGAGGTTCAATGAATATCC 424  
DB 32 GTCCAGATGAGACATCAGGAGGCTGAGTGAAGAGCCAGGCTTCCGTGAAGTCAAC 91  
OY 425 TGCAGAGCTTCTGTACTTCACTGAGCCACACCATGAGTGGTGAAGCAGACCAT 484  
DB 92 TGTCTGGCTCTGAGAGCAGCTTGGTAGTGTGAGCTGAGCTGGGTGAGACAGGCCCT 151  
OY 485 GGAAGAACTTGTGAGTGGATGAGTCTTAACTTCAATGAGTGAATCAATCAAC 544  
DB 152 GGACACAGCGCCCTGAGTGGAGAAATCATCCATGCTCATTCAGCAGCAACGCA 211  
OY 545 CAGAGTTCAGAGGAGGAGCCACATTTACTGAGACAGTCTCCAGCAGCAGCTACAG 604  
DB 212 CGAAGTTTCAGAGCAGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 271  
OY 605 GAGCTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 654  
DB 272 GAATGACAGCGCTGATATTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 321

RESULT 14  
AA345486 403 bp mRNA EST 21-APR-1997  
LOCUS EST51505 Gall bladder II Homo sapiens cDNA 5' end similar to  
DEFINITION similar to immunoglobulin kappa light chain, V region (GB:X06763),  
mRNA sequence.

ACCESSION AA345486  
NID 91997722  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; mitochondria; eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae;

REFERENCE 1 (bases 1 to 403)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fulcher,R.A.,  
Bull,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Maniatis,C.,  
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.,  
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,  
Kelley,J.M., Kelley,J.C., Liu,L.-T., Marmaros,S.M., Merrick,J.M.,  
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,J.,  
Small,K.V., Spriggs,T.A., Uteback,T.R., Weidman,J.F., Li,Y.,  
Bedarrik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dinke,D., Feng,D.-F., Ferlie,A., Fischer,C., Hastings,G.A.,  
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P., Olsen,H.,  
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,  
Fraser,C.M., and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence  
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)  
MEDLINE 96026280

COMMENT

Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3016699056  
Fax: 3016699423  
Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (<http://www.tigr.org/tdb/ngi/ngi.html>)  
Seq primer: M13 Reverse

FEATURES  
Location/Qualifiers  
1. 403

ORIGIN  
BASE COUNT 86 a 113 c 107 g 93 t 4 others  
ORIGIN

Query Match 15.6%; Score 110.8; DB 18; Length 403;  
Best Local Similarity 63.1%; Pred. No. 1.4e-24;  
Matches 188; Conservative 1; Mismatches 103; Indels 6; Gaps 1;

OY 44 CTCACGGGAGAGAGGTACACATGACTTGCAGGGGACCCCAAGTGT-----AAGTTAC 97  
DB 39 CTCACGGGAGAGAGAGGCGCCCTCTCTGAGGGGACGAGTGAAGAGAGAGTAC 98  
OY 98 ATGACCTGGTATCAGCAGAGAGGAGATCTCCCAAACTTGGATTATACCAATCC 157  
DB 99 TTAGCTGTGATACACAGAAACCTGCGCAGGCTCCAGGCTCTCATATGATGTCGTC 158  
OY 158 AACCTGCTTCTGAGAGTCCCTGCTGCTGAGTGGGCGGTGAGGAGCTTACTACT 217  
DB 159 AGGAGGCGCAGTGCATCCAGACAGTTCAGTGGAGTGGTGGAGCAGACTTCACT 218  
OY 218 CTCACAGTCAGCAGAGTGGAGGCTGAGATGCTGCTCAATTTACTGACAGAGTGAGT 277  
DB 219 CTCACATCAGCAGAGTGGAGGCTGAGATTTTGGAGGTATTTACTGACAGATAGT 278  
OY 278 GGTAGCCCAACCCAGCTTGGAGGAGGGGTCCAGAGCTGGAATTAAGGTTTCACTCTG 335  
DB 279 GACTCAGCTAGAGAGCTTGGCCCAAGGAGCAAGGTGAATCAACGAACGTGCTG 336

RESULT 15  
AA423447 398 bp mRNA EST 16-OCT-1997  
LOCUS ve08a03.r1 Soares mouse mammary gland NBMG Mus musculus cDNA clone  
DEFINITION 832492 5' similar to gb:M63438 IG KAPPA CHAIN PRECURSOR V-II  
REGION (HUMAN). gb:M35669 Mouse Ig aberrantly rearranged  
kappa chain mRNA V-02-C-region, (MOUSE); mRNA sequence.

ACCESSION AA423447  
NID 92102360  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 398)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theisling,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)  
COMMENT

Contact: Maira M/Mouse EST Project  
WashU-HHMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.wustl.edu  
This clone is available royalty-free through LBNL; contact the  
IMAGE Consortium (info@image.lbnl.gov) for further information.  
MG1:492708  
Seq primer: -28m13 rev2 ET from Amersham.

FEATURES  
source

Location/Qualifiers  
1..398

/organism="Mus musculus"  
/strain="C57BL/6J"  
/note="Organ: mammary gland; Vector: pT7p3D-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not I;  
Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5',  
TGTACCACTGAGAGGAGCGGCGCGGATGTTTCTTTTCTTTTCTTTT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT7/3 vector.  
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
constructed and normalized by Bento Soares and M.Fatima  
Bonaldo."

/db\_xref="taxon:10090"  
/clone="832492"  
/clone\_lib="Soares mouse mammary gland NDMG"  
/sex="male"  
/tissue\_type="mammary gland"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
BASE COUNT 101 a 100 c 102 g 95 t  
ORIGIN

Query Match 15.6%; Score 110.8; DB 24; Length 398;  
Best Local Similarity 63.2%; Pred. No. 1.4e-24;  
Matches 211; Conservative 0; Mismatches 107; Indels 16; Gaps 2;

QY 4 GACCTGACCTGACCCAGCTCTCCAGCAATCTGTCTGCACTCCAGGGAGAGGTACAC 63  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 66 GACATTGTGCTGACCAATCTCCAGCTTCTGCTGTCTCTAGGGCAG-AGGGCCAC 124  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 64 AATGACTGACAGGCGCCACCCCAAGTGA-----AGTTACATGCACTGGTA 108  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 125 CATCTCTGACAGAGCCAGCAAGAGTGTGTAATCTGSCATTAGTTTATGACTGGTT 184  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 109 TCAGCAGAGGAGGATCTCTCCCAACCTGATTTATACCATCCAACTGCTTC 168  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 185 CCAACAGAAACCAAGACAGGACCAACTCTCATCTATGCTGATCCCAACAAAGATC 244  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 169 TGGAGTCCCTGCTCTCTCAGTGGGCTGGTCTGGACCTTACTCTCTCAGAGTCAG 228  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 245 CGGGTCCCTGCGCAGTTAGTGGCAGTGGTCTGGACAGACTTCAGCTCAACATCCA 304  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 229 CAGAGTGAAGCTGAAGATGCTGCCATTATTACTGCGCAGAGTGAGTGTAGCCACC 288  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 305 TCCTATGGGAGGATGATCAATGATTTCTGTCAGCAAGTAAAGAGAGTTCCTCG 364  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 289 CACGTTGGAGGGGGTCCAAAGCTGGAATAAAA 322  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 365 GACGTTGGTGGAGGACCAAGTTGGAATCAAA 398  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: May 13, 1999, 09:40:43  
Job time: 361 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 09:35:42 ; Search time 59.56 Seconds

(without alignments)  
2248.983 Million cell updates/sec

Title: US-08-704-178-1

Perfect score: 712

Sequence: 1 ATGACCTGCAGCTGACCA.....GGACACGGTACCGTCTCC 712

Scoring table: IDENTITY\_NUC

Searched: 240622 seqs, 94065609 residues

Database : N\_Geneseq\_34.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	699.6	98.3	711	1	Single-chain anti-
2	696.4	97.8	711	1	Sequence encoding
3	694.8	97.6	711	1	Anti-erbB2 scfv cd
4	458	64.3	720	1	cDNA encoding an a
5	434	61.0	732	1	Fv(GP-2) immunosup
6	422	59.3	720	1	Sequence encoding
7	422	59.3	720	1	Single-chain anti-
8	356.2	50.2	913	1	Bispecific CD3-L6F
9	350.4	49.0	906	1	Anti-human SC sing
10	348.4	48.9	729	1	Fv(GP-4) immunosup
11	331.6	46.6	848	1	26-10 sFv coding s
12	320.6	45.0	738	1	OVB3 light and hea
13	319	44.8	797	1	CC49/212 SCA PLAP
14	317.8	44.6	1460	1	Sequence encoding
15	317.8	44.6	803	1	PLAP CC49/212 SCA
16	316.6	44.5	731	1	Sequence encoding
17	315.8	44.4	735	1	cDNA encoding an a
18	313.2	44.0	741	1	Coding sequence of
19	313.2	44.0	741	1	Single chain bindi
20	313.2	44.0	741	1	TR59, single chain
21	311.6	43.8	741	1	TR59, single poly
22	308.8	43.4	753	1	TR104b, Single po
23	308.8	43.4	753	1	Coding sequence of
24	308.8	43.4	753	1	Single chain bindi
25	308.8	43.4	753	1	TR104b single cha
26	298.6	41.9	724	1	18-2-3-/TR59, S
27	298.6	41.9	720	1	Coding sequence of
28	298.6	41.9	720	1	Single chain bindi
29	298.6	41.9	720	1	18-2-3-/TR59, sin
30	296.6	41.7	819	1	Sequence encoding
31	293	41.2	725	1	4-4-20 VL/ 217 / C
32	291.4	40.9	725	1	Sequence encoding
33	287.2	40.3	6799	1	Nucleic acid seque
34	285.8	40.1	819	1	Nematode salivary
35	285.8	40.1	819	1	Mus musculus antib
36	285.2	40.1	729	1	Coding sequence of
37	285.2	40.1	729	1	Single chain bindi
38	285.2	40.1	729	1	18-2-3-/TR59 singl
39	285.2	40.1	6727	1	Nucleic acid seque
40	284.2	39.9	1320	1	Nucleic acid seque
41	284.2	39.9	1299	1	Nucleic acid seque
42	282.8	39.7	836	1	Sequence of single
43	282.2	39.6	462	1	3H1 heavy chain va

ALIGNMENTS

44 282.2 39.6 462 1 T34542 Monoclonal anti-Id  
45 282.2 39.6 462 1 T99435 Anti-Idiotypic anti

RESULT	ID	Score	Length	DB	Match	Score	Length	ID	Description
1	T65006	98.3%	711	1	Single-chain anti-erbB2 antibody e23(Fv) cDNA.				
		Best Local Similarity 99.9%; Pred. No. 4,7e-177;			Single-chain antibody; variable region; light chain; heavy chain;				
		Matches 711; Conservative 0; Mismatches 0; Indels 1; Gaps			breast cancer; ovarian cancer; non-small cell lung carcinoma;				
					immunodiagnosis; treatment; cytotoxic agent; erbB-2, ds.				
					Mus musculus.				
					Synthetic.				
					Key				
					Location/Qualifiers				
					1.711				
					/tag= a				
					/product= e23(Fv)				
					US587458-A.				
					24-DEC-1996.				
					07-OCT-1991; 772270.				
					07-OCT-1991; US-772270.				
					30-JUN-1992; US-906555.				
					14-MAY-1993; US-061092.				
					(ARON-) ARONEX PHARM INC.				
					Bird RE, Kasprzyk PG, King CR;				
					WPI; 97-064831/06.				
					P-PSDB; WI18185.				
					Single chain antibodies specific for erbB-2 protein, gp185 - with				
					labels or cytotoxin, useful for detection and treatment of tumour				
					cells expressing this protein				
					Example 8: Columns 25-28; 28pp; English.				
					The present cDNA sequence codes for a claimed single-chain antibody,				
					designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23				
					was generated by immunising mice with M/erbB-2 cells overexpressing				
					the gp185 protein, removing spleen cells and producing hybridomas				
					by standard techniques. Messenger RNA coding for the anti-erbB-2				
					monoclonal antibody was isolated and converted to cDNA. Regions				
					coding for the heavy- and light- chain variable regions were then				
					amplified by PCR and joined via a sequence encoding a peptide				
					linker. The resulting single-chain antibody is useful for in vitro				
					diagnosis of tumour cells which overexpress the erbB-2 gp185				
					marker, e.g. breast, ovarian and non-small cell lung carcinomas,				
					and, when coupled to a cytotoxic agent, to treat such tumours.				
					Sequence 711 BP; 175 A; 182 C; 189 G; 164 T;				
					SO				

Db	240	TGAAGATGCTCCCACTTATTACTGCCAGCAGGTGAGTGTAGGCCACCACCCAGTTGGAGG	29
Qy	301	GGGGTCCAGCTGGGAATTAAGTGTACTCTCGTGTTCGTAAATCTCTGAAGTAA	360
Db	300	GGGGTCCAGCTGGGAATTAAGTGTACTCTCGTGTTCGTAAATCTCTGAAGTAA	359
Qy	361	AGGTGTGACGTGCAGAGAGTCAGACCTGAGGTGGTAAAGCTTGAGTTCATGAAGAT	420
Db	360	AGGTGTGACGTGCAGAGAGTCAGACCTGAGGTGGTAAAGCTTGAGTTCATGAAGAT	419
Qy	421	ATCTGCAAGACTTCTGGTTACTCTATTCTACTGCCCACACCATGAATGGGTGAAGCAGAG	480
Db	420	ATCTGCAAGACTTCTGGTTACTCTATTCTACTGCCCACACCATGAATGGGTGAAGCAGAG	479
Qy	481	CCATGGAAGAAGACCTTGTAGTGGATTGGACATTAAATCCCTTACAAATGGTAACTACTA	540
Db	480	CCATGGAAGAAGACCTTGTAGTGGATTGGACATTAAATCCCTTACAAATGGTAACTACTA	539
Qy	541	CAACAGAGAATTTCAAGGGCAAGGCCACATTACTGTAGACAAAGTCCTCCAGCAGCCTA	600
Db	540	CAACAGAGAATTTCAAGGGCAAGGCCACATTACTGTAGACAAAGTCCTCCAGCAGCCTA	599
Qy	601	CATGAGACTCTCAAGTCTGACATTTGAGGACTCTGCACTATTACTGTGCAGAGAGGT	660
Db	600	CATGAGACTCTCAAGTCTGACATTTGAGGACTCTGCACTATTACTGTGCAGAGAGGT	659
Qy	661	TACGAGACTGATCTCGATGCTCTGGGGGCGAGGGAGACACGGTCAACGCTTCC	712
Db	660	TACGAGACTGATCTCGATGCTCTGGGGGCGAGGGAGACACGGTCAACGCTTCC	711

RESULT	2
Q55180	
ID	Q55180 standard; cDNA; 711 BP.

DT	21-JUL-1994	(first entry)
DV	Sequence encoding the single chain anti-erbB antibody, Ab no.23.	
KW	Single chain anti-erbB antibody; cancer therapy; prevention;	
KW	monoclonal antibody; ss.	
OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	cds	1..171
PT	WO9400136-A.	/+tag- a
PD	06-JAN-1994.	
PF	21-OCT-1992; U08545.	
PR	30-JUN-1992; US-906555.	
PA	(MOLE-) MOLECULAR ONCOLOGY INC.	
PI	Kasprzyk PG, King CR:	
DR	WPI: 94-025878/03.	
PT	P-PSDB: R45442.	
FT	Treatment of malignancies over-expressing ERB-B2 - using at least	
PT	2 monoclonal antibodies which recognise different epitopes on	
PT	gp185	
PS	Example: Fig 7: 37PD: English.	
CC	The source of human erbB-2 protein for the prodn. of antibodies no	
CC	23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell	
CC	engineered to express the human erbB-2 protein on its surface (N/	
CC	erbB-2). Abs no. 21 and 23 are directed against the extracellular	
CC	domain of gp185 erbB-2. Nude mice manipulated to produce rapidly	
CC	growing tumours were used in a trial of the efficacy of the Abs. In	
CC	animals given a combination of the 2 Abs, tumours completely	
CC	regressed after 11 days.	
QC	Sequence 711 BP; 175 A; 183 C; 187 G; 165 T;	

Query Match	97.8%	Score 696.4	DB 1	Length 711
Best Local Similarity	99.6%	Pred. No. 3.3e-176		
Matches 709	Conservative	0	Mismatches 2	Indels 1
				Gaps 1
OY	1	ATGACACCTGCACCTATACCCACCATCTCTCCAGCAATCTCTGCATCTCCAGGGAGAAAGTGA	60	
		ATGGACCTGCACCTATACCCACCATCTCTCCAGCAATCTCTGCATCTCCAGGGAGAAAGG-1	60	

QY	61	CACATACACTGCAAGGGCCACCCCAAGTAAATTACATCATCTGGATACAGAGAACC	120
Db	60	CACATAGCTTGCAAGGGCCACCCCAAGTAAATTACATCATCTGGATACAGAGAACC	119
QY	121	AGATCCTCCCCCAACCTTGATTTATACACATCCACATCTGCTTGGATCCCTGC	180
Db	120	AGGATCCTCCCCCAACCTTGATTTATACACATCCACATCTGCTTGGATCCCTGC	179
QY	181	TCGCTTAGTGGGGGTGGGTCTGGAGACCTTACTCTCACAAGTACAGAGATGGAGGC	240
Db	180	TCGCTTAGTGGGGGTGGGTCTGGAGACCTTACTCTCTACAGTACAGAGATGGAGGC	239
QY	241	TGAAGATGCTGCCACTTATTACTGCCAGAGTGGAGTCTGAGCCACCCAGTTGGAGG	300
Db	240	TGAAGATGCTGCCACTTATTACTGCCAGAGTGGAGTCTGAGCCACCCAGTTGGAGG	299
QY	301	GGGGTCCAAAGCTGGAAATTAAGGTTTACTCTGCTTCTGGTAAATCTTCTGAAGTAA	360
Db	300	GGGGTCCAAAGCTGGAAATTAAGGTTTACTCTCTGGTAAATCTTCTGAAGTAA	359
QY	361	AGGTGTCCAGCTGCAGAGAGTCAGAGCTGAGTGGTAAACCTCGAAGGTTCAATGAAGAT	420
Db	360	AGGTGTCCAGCTGCAGAGATTCAGAGCTGAGTGGTAAACCTCGAAGGTTCAATGAAGAT	419
QY	421	ATCTGTCAAGACTTCTGGTTACTATTCACTGCGCCACACCATGAATGGGTGAAGCAGAG	480
Db	420	ATCTGTCAAGACTTCTGGTTACTATTCACTGCGCCACACCATGAAGTGGGTGAAGCAGAG	479
QY	481	CCATGGAAGAAGCACTTAGTGGATTTGACTATTAAACCTTAAGAAGTATACTAACTA	540
Db	480	CCATGGAAGAAGCACTTAGTGGATTTGACTATTAAACCTTAAGAAGTATACTAACTA	539
QY	541	CAACACAGAAGTTCAAAGGGCAAGGCCAATTTACTGTGACAAGTCTGCCAGCAGACCTTA	600
Db	540	CAACACAGAAGTTCAAAGGGCAAGGCCAATTTACTGTGACAAGTCTGCCAGCAGACCTTA	599
QY	601	CATGAGAGTCTCTCAGTTGACATCTGAGCACTCTGCAGTCTATTACTGTGCAAGAGAGGT	660
Db	600	CATGAGAGTCTCTCAGTTGACATCTGAGCACTCTGCAGTCTATTACTGTGCAAGAGAGGT	659
QY	661	TACGAGACGTGACTTGATGTCTGGGGCGAAGGAGACCAAGGTCACGCTTCC	712
Db	660	TACGAGACGTGACTTGATGTCTGGGGCGAAGGAGACCAAGGTCACGCTTCC	711

### RESULT 3

ID T17728 standard; CDNA; 711 BP.  
 AC T17728;  
 DT 21-MAY-1996 (first entry)  
 DE Anti-erbB2 scFv CDNA.  
 KW Oncoprotein; erbB2; cell proliferation; tumour; cancer;  
 KW Intracellular antibody homologue; single chain antibody; scFv;  
 KW gene therapy; ds.  
 OS Synthetic  
 PN W09607321-A1.  
 PD 14-MAR-1996.  
 PF 23-SEP-1995; 010740.  
 PR 06-SEP-1994; US-301339.  
 PR 06-JUN-1995; US-468252.  
 PA (UABR-) UAB RES FOUND.  
 PI Curjel DT, Deshane J;  
 PI WPI; 96-171307/17.  
 DR P-PSDB; R94020.  
 PT Inhibition of proliferation or survival of, esp. malignant erbB2,  
 PT cells - by introducing nucleic acid mol. encoding antibody homologue  
 PT which is expressed and binds, pref. erbB2, protein intracellularly  
 PS Claim 42; Page 29-30; 48pp; English.  
 PC A nucleic acid comprises a first sequence encoding a signal peptide  
 CC (R94019) linked to a second sequence (T17728) encoding a single  
 CC chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.  
 CC The anti-erbB2 scFv portion is obtained by PCR using e23scv



QY 657 - -GGGTACAGGACTGCTGATGCTGCTGGGGGCGAGGAGCAGGTCACGCTGCC 712  
 Db 660 CTATGTATACGGCAGCTTTGACTACTGGGGCCAGGACACACTCTACAGCTGCC 717

## RESULT 5

Q73678 ID Q73678 standard; DNA; 732 BP.  
 AC Q73678;  
 DT 09-MAY-1995 (first entry)  
 DE Fv(GP-2) immunosuppressive.  
 KM MAb; monoclonal antibody; hybridoma; interleukin-2; IL-2;  
 KW Fv; antibody variable region; GP-2; Fv(GP-2);  
 OS Immunosuppressive; ss.  
 PN EP-621338-A.  
 PD 26-OCT-1994.  
 PF 21-APR-1994; 106257.  
 PR 21-APR-1993; JP-094491.  
 PA (AJIN) AJINOMOTO KK.  
 PI Hamura Y, Kanayama Y, Nakazawa H, Shimamura T, Sugamura K;  
 PI Takeishi T;  
 DR WPI; 94-325948/41.  
 DR P-PSDB; R60780.  
 PT Immunosuppressant polypeptide - has ability to block  
 PT Interleukin-2 response  
 PS Claim 17: Page 29; 37pp; English.  
 CC MAb capable of binding to the gamma chain of the IL-2 receptor, and  
 CC thus of blocking the IL-2 response, is produced by mouse hybridoma  
 CC line GP-2 (FERM BP-4641). DNA encoding the variable region of  
 CC this MAb was expressed in E. coli, yielding Fv(GP-2) with  
 CC immunosuppressive activity.  
 SQ Sequence 732 BP; 180 A; 204 C; 182 G; 166 T;

Query Match 61.0%; Score 434; DB 1; Length 732;  
 Best Local Similarity 78.6%; Pred. No. 1.5e-106;  
 Matches 574; Conservative 1; Mismatches 136; Indels 19; Gaps 4;

QY 1 ATGACCTGACGCTGACCCAGCATCTGCTGCTGATCTCCAGGGAGAAAGTA 60  
 Db 1 ATGGACATCTCGCTGACCCAGCATCTGCTGCTGATCTCCAGGGAGAAAGTA 59  
 QY 61 CACATGACTGTCAGGCGACCCCAAGTGA-----AGTACATGACCTGTAACGCA 114  
 Db 60 CACATGACTGTCAGGCGACCTGACACTCAAGTGAAGTCCAGTACTGCACTGTAACGCA 119  
 QY 115 GAAGCCAGAGATCTCCCAAACTTGATTTATACCAATCAACCTKCTCTGGAGT 174  
 Db 120 GAAGCCAGAGATCTCCCAAACTTGATTTATAGCACATCAACCTGCTCTGGAGT 179  
 QY 175 CCCCTGCTGCTTACAGTGGGGTGGGCTGAGACCTCTTACTCTCTACAGTCACAGAGT 234  
 Db 180 CCCAGCTGCTTACAGTGGGGTGGGCTGAGACCTCTTACTCTCTACAAACACACAGAT 239  
 QY 235 GGAGCTGGAAGATCTGCGACCTATTACTGCGACAGTGAAGTGGTGAAGCCACAGCTT 294  
 Db 240 GGAGCTGGAAGATCTGCGACCTATTACTGCGACAGTGAAGTGGTGGTGGCTGAGCTT 299  
 QY 295 CGAGGGGGGTCAGAGCTGGAATAAAAGTTCTA---CCTCTGTTCTGGTAAATCTTC 351  
 Db 300 CGGTGCTGGGACCAAGCTGAGGCTCAAAAGTGAACAATCCTCAGAGATCTGCTCCGATC 359  
 QY 352 TGAAGGTAAAGGTGTCAGCTGCAAGAGTCAAGACTGAGGTGGTGAAGCTGAGGCTTC 411  
 Db 360 CAAAGGACCGAGGCTCAAACTCGAGAGTCTGAGCTGGTGGGCTGAGAGCTTC 419  
 QY 412 AATGAAGATCTGCAAGATCTTGTTACTTACTTCTGCTGCTGCGACACAGATGAAGTGGGT 471  
 Db 420 AGTGAAGCTGCTGCAAGGCTTCTGCTTACACATTCACAGCTACTGAGATGCACTGGGT 479

QY 472 GAAGCAGACCCATGAGAAAGACCTTGAGTGGATGATGATTAATCTTACATGGTGA 531  
 Db 480 GAAGCAGACCCATGAGAAAGACCTTGAGTGGATGATGATTAATCTTACATGGTGA 539  
 QY 532 TACTTAAGTAAACAGAGTTCAGAGGCGACACTTACGTAGTACAGAGTCTCCAG 591  
 Db 540 TACTTAAGTAAACAGAGTTCAGAGGCGACACTTACGTAGTACAGAGTCTCCAG 599  
 QY 592 CACAGCTTACATGAGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATG 651  
 Db 600 CACAGCTTACATGAGAGTCTCTGATGATGATGATGATGATGATGATGATGATGATG 659  
 QY 652 AAGGAGGTTTACGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 702  
 Db 660 AAGGAGGTTTACGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 719  
 QY 703 CACCGTCTCC 712  
 Db 720 CACCGTCTCC 729

## RESULT 6

Q55181 ID Q55181 standard; cDNA; 720 BP.  
 AC Q55181;  
 DT 21-JUL-1994 (first entry)  
 DE Sequence encoding the single chain anti-erbB2 antibody, Ab no. 21.  
 KM Single chain anti-erbB1 antibody; cancer therapy; prevention;  
 KW monoclonal antibody; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT cds 1..171  
 FT CDS /tag- a  
 PN M09400136-A.  
 PD 06-JAN-1994.  
 PF 21-OCT-1992; U08545.  
 PR 30-JUN-1992; US-906555.  
 PA (MOLE-) MOLECULAR ONCOLOGY INC.  
 PI Kasprzyk PG, King CR;  
 DR WPI; 94-025878/03.  
 DR P-PSDB; R45443.  
 PT Treatment of malignancies over-expressing ERB-2 - using at least  
 PT 2 monoclonal antibodies which recognise different epitopes on  
 PT gp185  
 PS Example; Fig 8; 37pp; English.  
 CC The source of human erbB-2 protein for the prodn. of antibodies no.  
 CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell  
 CC engineered to express the human erbB-2 protein on its surface (N/  
 CC erbB-2). Abs no. 21 and 23 are directed against the extracellular  
 CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly  
 CC growing tumours were used in a trial of the efficacy of the Abs. In  
 CC animals given a combination of the 2 Abs, tumours completely  
 CC regressed after 11 days  
 SQ Sequence 720 BP; 184 A; 178 C; 179 G; 179 T;

Query Match 59.3%; Score 422; DB 1; Length 720;  
 Best Local Similarity 79.3%; Pred. No. 2.3e-103;  
 Matches 511; Conservative 1; Mismatches 131; Indels 1; Gaps 1;

QY 8 TGCAGCTGACCCAGTCTCAGCAATCTGCTGCTGATCTCCAGGGAGAGATACATG 67  
 Db 2 TGCAGCTGACCCAGTCTCAGCAATCTGCTGCTGATCTCCAGGGAGAGAGATACATG 60  
 QY 68 ACTTGACGGGCGACCCCAATGTAAGTATGATGATGATGATGATGATGATGATGATG 127  
 Db 61 ACTTGACGGGCGACCCCAATGTAAGTATGATGATGATGATGATGATGATGATGATG 120  
 QY 128 TCCCCAAACCTTGATTTATACCATCCAACTKCTCTGAGAGTCCGCTGCTTC 187  
 Db 121 TCCCCAAACCTTGATTTATGACACATCCAACTGCTCTGAGAGTCCGCTGCTTC 180  
 QY 188 AGTGGCGGTGGTCTGGACCTCTTACTCTCTCAGAGTGAAGTGAAGTGAAGTGAAGT 247

Db 181 ACTGTCAGTGGGCTGGAACACTTACTCTCAGCATCAGACGATGAGGCTGAAGT 240  
QY 248 GCTGCTACTTATTACGCGACGATGAGTCTAGCCACCCACGTTCCGAGGGGGTCC 307  
Db 241 GCTGCTACTTATTATGATCAGGGGAGTGGTACCCATTACGTTGGGCTCGGGGACA 300  
QY 308 AAGCTGGAATATAAGGTTCTACCTGTGTTGTGTAATCTTCTGAGAGTAAAGGTGTG 367  
Db 301 AAGTTGGAAATATAAGGTTCTACCTCGGATCTGTAATCTTCTGAGAGTAAAGGTGTG 360  
QY 368 CAGTCGACGAGATCAGGACCTGAGGCTGGAAGCTGGAGTCAATGAATATCCCTCC 427  
Db 361 CAGTCGACGAGATCTGGGGTTAGCTGTCCAGAGGGGGCTTATGTCAGTTGTCTCC 420  
QY 428 AAGACTTCTGTTACTCATCTACGCGCCACACCATGATGAGGTGGAGAGCCATGGA 487  
Db 421 AAGCTTCTGATTCACACTTAAAGACTTATATCCACTGGGTGAGAGAGGCTGAA 480  
QY 488 AAGAACTTGAAGTGAATGAGCTTATTAATCCTTACATATGATGATACATAACACG 547  
Db 481 CAGGGCTGGAATGATGATGATGATTCATCCTGAGAAATGATATCTATATGACCCG 540  
QY 548 AAGTTCAGGAGGAGGACCATTTACTCTAGACAACTGCTCCAGACACACCTACATGAG 607  
Db 541 AATTCAGAGGAGGAGGACCATTTACTCTAGACAACTGCTCCAGACACGCGCTACCTTCAG 600  
QY 608 CTCTCAGTCTGACATCTGAGACTCTGAGACTTATTACTGTGC 651  
Db 601 CTCAGACGCTGACATCTGAGGACACTGCCGTCTATTACTGTGC 644

## RESULT 7

T65007 standard; cDNA; 720 BP.

AC T65007; 1997 (first entry)  
DE Single-chain anti-erbB2 antibody e21(Fv) cDNA.  
KW Single chain antibody; variable region; light chain; heavy chain;  
breast cancer; ovarian cancer; non-small cell lung carcinoma;  
immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.  
OS Mus musculus.  
FH Synthetic.  
FH Key Location/Qualifiers  
FT mat\_peptide 1..720  
FT /\*tag= a  
FT /\*product= e21(Fv)  
FN US5587458-A.  
PD 24-DEC-1996.  
PF 07-OCT-1991; 772270.  
PR 07-OCT-1991; US-772270.  
PR 30-JUN-1992; US-906555.  
PR 14-MAY-1993; US-061092.  
PA (ARON-) ARONEX PHARM INC.  
PI Bird RE, Kasprzyk Pg, King CR;  
DR WPI; 97-064831/06.  
DR P-PsDB; W15186.  
PT Single chain antibodies specific for erbB-2 protein. gp185 - with  
cells expressing this protein  
PS Example 9; Columns 27-30; 28pp; English.  
CC The present cDNA sequence codes for a claimed single-chain antibody,  
designated e21(Fv), which binds to erbB-2. Monoclonal antibody e21  
was generated by immunising mice with N/erbB-2 cells overexpressing  
the gp185 protein, removing spleen cells and producing hybridomas  
by standard techniques. Messenger RNA coding for the anti-erbB-2  
monoclonal antibody was isolated and converted to cDNA. Regions  
coding for the heavy- and light- chain variable regions were then  
amplified by PCR and joined via a sequence encoding a peptide  
linker. The resulting single-chain antibody is useful for in vitro  
diagnosis of tumour cells which overexpress the erbB-2 gp185  
marker, e.g. breast, ovarian and non-small cell lung carcinomas,  
and, when coupled to a cytotoxic agent, to treat such tumours.

SQ Sequence 720 BP; 184 A; 178 C; 179 G; 179 T;

Query Match 59.3%; Score 422; DB 1; Length 720;  
Best Local Similarity 79.3%; Pred. No. 2,3e-103;  
Matches 511; Conservative 1; Mismatches 131; Indels 1; Gaps 1;

QY 8 TGCAGCTGACCCAGTCTCCAGCAATCTGTGTGATCTCCAGGGGAGAGATGACAAATG 67  
Db 2 TGCAGCTGACCCAGTCTCCAGCAATGATGTGATCTCCAGGGGAGAGATGACAAATG 60  
QY 68 ACTTCAGGGCCACCCCAAGTATGATACATGCACTGGTATACACAGAGAGGCTAGATCC 127  
Db 61 ACTTCAGGGCCAGTCTCAAGTATGATACATGCACTGGTATACACAGAGAGGCTAGATCC 120  
QY 128 TCCCCCAAACTTGGATTATATACATCAACCTTGGTGTGAGTCCCTGCTGCTTC 187  
Db 121 TCCCCCAAACTTGGATTATATACATCAACCTTGGTGTGAGTCCCTGCTGCTTC 180  
QY 188 ACTGCGGTGGGTCTGGGACCTTACTCTCTCAGAGTCCAGAGAGAGGCTGAGAT 247  
Db 181 ACTGCGAGTGGGTCTGGGAACTTACTCTCTCAGAGTCCAGAGAGAGGCTGAGAT 240  
QY 248 GCTGCTACTTATTACTCCAGCAGTGAAGTCTAGCCACCCAGCTTGGAGGGGGTCC 307  
Db 241 GCTGCTACTTATTATTATGATGATGAGGGAGTGGTACCACTTACGTTGGCTCGGGGACA 300  
QY 308 AAGCTGGAATATAAGGTTCTACCTCTGTTGTTGTAATCTTCTGAGAGTAAAGGTGTG 367  
Db 301 AAGTTGGAAATATAAGGTTCTACCTCGGATCTGTAATCTTCTGAGAGTAAAGGTGTG 360  
QY 368 CAGTCGACGAGATCAGGACCTGAGGTGATGAAAGCTGGAAGTTCAATGAATATCTGTC 427  
Db 361 CAGTCGACGAGATCTGGGGTTAGCTGTGCCAGAGGGGGCTTATGTCAGTTGTCTCC 420  
QY 428 AAGACTTCTGTTACTCATCTACGCGCCACACATGAACTGGGTGAAGAGCCATGGA 487  
Db 421 AAGCTTCTGATTCACACTTAAAGACTTATATATCCACTGGGTGAAGAGGCTGTA 480  
QY 488 AAGAACTTGAAGTGAATGAGCTTATTAATCCTTCAATGATGATACATAACACG 547  
Db 481 CAGGGCTGGAATGATGATGATGATTCATCCTGAGAAATGATATGATATGATACCCG 540  
QY 548 AAGTTCAGGAGGAGGACCATTTACTCTAGACAACTGCTCCAGACACGCTACATGAG 607  
Db 541 AATTCAGAGGAGGAGGACCATTTACTCTAGACAACTGCTCCAGACACGCGCTACCTTCAG 600  
QY 608 CTCTCAGTCTGACATCTGAGACTCTGAGACTTATTACTGTGC 651  
Db 601 CTCAGACGCTGACATCTGAGGACACTGCCGTCTATTACTGTGC 644

RESULT 8  
O81076 standard; cDNA; 913 BP.  
ID O81076;  
AC O81076;  
DE 14-MAR-1995 (first entry)  
KW Bispecific CD3-J6Fvlg antibody derivative coding sequence.  
KW fusion protein; recombinant bispecific single chain antibody;  
helical peptide linker; anti-L6 antibody; tumour cell antigen;  
anti-CD3 antibody; variable region; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT signal\_peptide 7..75  
FT /\*tag= a  
FT /\*note= "L6 light chain variable region leader"  
FT cds 7..912  
FT /\*tag= b  
FT /\*tag= 76..819  
FT misc\_rna  
FT /\*tag= c  
FT /\*note= "encodes CD3 VL-VH fusion"  
FT /\*tag= 820..828  
FT misc\_rna  
FT /\*tag= d

FT misc\_rna /note="encodes hinge region of Fc domain"  
 FT 829..912 /tag="e"  
 FT /note="encodes helical peptide linker"  
 PN EP-610046-A.  
 PD 10-AUG-1994.  
 PF 31-JAN-1994: 300692.  
 PR 01-FEB-1993: US-013420.  
 PR 13-SEP-1993: US-121054.  
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.  
 PI Bajorath J, Fell PH, Gilliland LK, Hayden MS, Ledbetter JA,  
 PI Linsley PS;  
 DR WPI: 94-250885/31.  
 DR P-PSDB: R60206.  
 PT Expression vector encoding bispecific fusion protein - having  
 PT binding domains for separate targets joined by helical peptide,  
 PT useful e.g. for diagnosis and treatment  
 PS Example 1: Fig 11: 50pp: English.  
 CC The VL and VH sequences of the anti-CD3 hybridoma G19-4 were  
 CC amplified by PCR methods. A gene fusion was constructed from the  
 CC two amplified domains and a (Gly4Ser)3 linker. The amino  
 CC termini of the VL-VH fusion cassette was fused at the SalI site to  
 CC the L6 light chain variable region leader peptide and the  
 CC carboxy-terminus was fused directly to the hinge region of the Fc  
 CC domain at the BclI site and/or to a short "helical peptide linker  
 CC to construct the bispecific CD3-L6vlg antibody derivative. The  
 CC variable regions for L6 were fused in frame to the opposite end of  
 CC the helical linker (not included in 081076).  
 SQ Sequence 913 BP: 252 A: 223 C: 228 G: 210 T:

Query Match 50.0%; Score 356.2; DB 1; Length 913;  
 Best Local Similarity 71.9%; Pred. No. 7.2e-86;  
 Matches 537; Conservative 0; Mismatches 173; Indels 37; Gaps 4;

QY 2 TGGAGCTGAGCTGAGCCAGCTCCAGCAATCCTGTGTCGATCCAGGGAGAGCTAC 61  
 DB 74 TCACATCCAGATGACACAGACACTACCTCCTGCTGCTGGGAGACAGT-C 132  
 QY 62 ACAATGACTGAGGAGCCACCC--CAAGTGTAACTTACATGCATGATACAGAGAAG 118  
 DB 133 ACATAGATGAGGAGGAGTGCAGACATTCGCAATTATTAACTGATACGAGAGAA 192  
 QY 119 CCAGATGCTCCCAACCTTGATTTATACCAATCCAACTKCTCTGATGCTCCT 178  
 DB 193 CCAGATGAGACTGTAACTCCGATCTACTACATCAAGATTACACACAGAGAGTCCA 252  
 QY 179 GCTCGTTAGTGGGGGTGGTGTGGACTCTTACTCTCTACATCAGTACAGAGAGAG 238  
 DB 253 TCAAGGTAGTGGAGTGGTGTGGAGATTAATTCCTCCATTCGCAACCTGCGAA 312  
 QY 239 GGTGAGATGCTGCACTTATTTACTGCCAGAGTGGAGTGTGGCCACCCAGCTCGGA 298  
 DB 313 CCAGAAATATTTGCCACTTACTTTGCCAAGAGGTAATACGCTTCGGTGAGCTTCGG 372  
 QY 299 GGGGGTCCAGAGCTGGAATTAAGGTTTACCTGCTGTGTAATCTTCTGAAGGT 358  
 DB 373 GGGAGCACCACCACTGTAACTAAGGAGAGTGGTGGGCTCGGGAGGTGGTGG 432  
 QY 359 AAAGT-----GTGCAGCTCAGAGAGTCCAGAGCTAGAGCTGAGGTGGT 397  
 DB 433 TCGGGTGGGGGGGATCTATCGATGAGGTCCACCTGACAGCTGAGCTGAAGTGG 492  
 QY 398 AAGCTCGAGGTTCATGAGATATCTTCGAAAGATTCCTGTTACTCATTCAGTGGCCAG 457  
 DB 493 AAGCTCGAGGTTCATGAGATATCTTCGAAAGATTCCTGTTACTCATTCAGTGGCCAG 552  
 QY 458 ACCATGAGTGGTGAAGAGAGCCATGAGAAAGACCTTGAGTGGATTTAAT 517  
 DB 553 ATCGTGAATGGCTGAAGAGAGCCATGAGAAAGACCTTGAGTGGATTTAAT 612  
 QY 518 CCTTAAATGGTATCTACTACTCAACAGAGTTCAGAGGAGCCACATTTACTGTA 577

DB 613 CCATCAAAAGTCTTACTACTACCAACCAAAATTCAGAGGCCAGCCATTAATCTGA 672  
 QY 578 GACAAGTCTGCAGACACAGCTTACATGAGAGTCTCCATGACATCTGAGACTCTGCA 637  
 DB 673 GACATGATTCACAGACACAGCTTACATGAGAGTCTCCATGACATCTGAGACTCTGCA 732  
 QY 638 GTCTATTTACTGTGCAAG-----GAGGTTACGAGTGTGTTACTTCAATGTCTGG 685  
 DB 733 GTCTATTTACTGTGCAAGAGTGTGTTACTTCAATGTCTGG 792  
 QY 686 GCGCAGGAGCAGACAGTCCAGCTCTCC 712  
 DB 793 GCGCAGGAGCAGACAGTCCAGCTCTCC 819

RESULT 9  
 V00611  
 ID V00611 standard; DNA: 906 BP.  
 AC 14-JUL-1998 (first entry)  
 DE Anti-human SC single chain Fv/protamine fusion protein encoding DNA.  
 KW Fusion protein; SC single chain Fv/protamine fusion protein; SECR;  
 KW exogenous gene; serpin enzyme complex receptor; gene therapy; target;  
 KW binding moiety; ds.  
 OS Homo sapiens.  
 OS Mus sp.  
 FH Key  
 FT CDS Location/Qualifiers  
 FT 1..906  
 FT /product="anti-human SC single chain Fv/protamine  
 FT fusion protein"

PN W09746100-A1.  
 PD 11-DEC-1997.  
 PE 03-JUN-1997: U09858.  
 PR 03-JUN-1996: US-656906.  
 PA (UYCA-) UNIV CASE WESTERN RESERVE.  
 PI Davis PB, Ferkol TW, Zlady A;  
 DR WPI: 98-041783/04.  
 DR P-PSDB: W37085.  
 PT Delivering compacted exogenous nucleic acid to cells - by targeting  
 PT the serpin enzyme complex receptor, used in gene therapy  
 PS Example 9: Pages 120-121: 158pp: English.  
 CC This DNA encodes an anti-human SC single chain Fv/protamine fusion  
 CC protein containing a target binding moiety capable of binding to a  
 CC serpin enzyme complex receptor (SECR), and a nucleic acid binding moiety.  
 CC This can be used in a method for delivering an oligonucleotide to a  
 CC mammalian cell. The method comprises conjugating the target binding  
 CC moiety to a nucleic acid binding moiety to form a carrier and coupling  
 CC the carrier to an expression vector encoding one or more gene products  
 CC to form a pharmaceutical composition. A mammalian cell having on its  
 CC surface SECR, is contacted with the pharmaceutical composition under  
 CC conditions allowing binding to the receptor resulting in delivery of the  
 CC pharmaceutical composition to the interior of the cell. The composition  
 CC and method are used for the introduction of exogenous genetic material  
 CC into target host cells expressing SECR on their surface. The nucleic acid  
 CC may encode a functional wild-type or mutant gene or may be an antisense  
 CC sequence or other nucleic acid having a therapeutic effect. The fusion  
 CC sequence may comprise a protein portion having therapeutic properties,  
 CC e.g. enzymatic activity, cytokine activity and antibiotic activity which  
 CC is delivered to a cell surface via the SECR binding moiety. The nucleic  
 CC acid can be compacted at high concentrations with the carrier molecule at  
 CC a critical salt concentration. The condensation of such complexes  
 CC provides structural features to the DNA/cationic lipid complex that  
 CC prolong in vivo expression.  
 SQ Sequence 906 BP: 226 A: 231 C: 257 G: 192 T:

Query Match 49.2%; Score 350.4; DB 1; Length 906;  
 Best Local Similarity 71.0%; Pred. No. 2.5e-84;  
 Matches 514; Conservative 0; Mismatches 191; Indels 19; Gaps 3;  
 QY 1 ATGAGCTGACACTGACCCAGTCTCCAGCAATCCTGTGCTGATCTCCAGGAGAGGTA 60

```
Db 1 ATGACATGTGCTGACCGAGTCTCCAGCTTCTTGGCTGTGCTCTAGGGCAGAGGGC- 60
QY 61 CACAATGACTTCAGAGGCCACCCCAAGTGA-----AGTTACATGCACTG 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 CACCATCTCTGAGAGCCAGCCGAAAGTGTGATTAATTAATGATAGCTAGTTTAACTG 119
QY 106 GTATTCAGCAAGAGCCAGATCTCCCAACCTTGATTTATACCAATCCAACTKGC 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 GTTCCACACAAACACAGAGAGCCCAACCTTCATATATGCTCAATCCAAACAGG 179
QY 166 TTCTGAGTCCCTGCTGCTTCACTAGTGTGGGTGGGTGGAGCCCTTACTCTGACAGT 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 ATCCGGGGTCCCTGCGAGCTTGTAGTGGTGTGGGACAGCTTCAAGCTTCAACAT 239
QY 226 CACAGAGTGGAGGCTGAGATGCTGCACTTATTAATGAGAGTGGAGTGTAGCCC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 CCATCTATGAGAGAGATGATATCTGCAATGTATTTCTGACAGCAAGTAAAGGGGTTC 299
QY 286 ACCAGCTTGGAGGGGGGTCCAAAGCTGGAAATAA---AGTTCTCTCTGTGTTCTGG 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 GTACAGCTTGGAGGGGGGCAAAAGCTGGAGATAAAGAGGGGGGTCTCGGGCGGTGG 359
QY 343 TAAATCTTGAAGGTAAAGGTGTGACAGTGTGACAGAGTCAAGAGTGGAGTGGAGCC 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 CGGATCCGGGGGGGCTCTGAGTGTGACAGTGTGACAGTGTGACAGTGTGAGAGCC 419
QY 403 TGGAGTTCATGAAGATATCTGCAAGAGTCTGTGTTACTTCACTGAGCCACACAT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 TGGGGCTTCAAGTATCTCTCAAGAGTCTGTGATCAACATTAATGAATACAT 479
QY 463 GAATCTGGTGAAGAGCCATGTAAGAAACCTTGATGTGATTTGATTTAACTCTTA 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCATCTGGTGAAGAGCCATGTAAGAAACCTTGATGTGATTTGATTTAACTCTTA 539
QY 523 CAATGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 CAATGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 599
QY 583 GTGTCACAGACAGCTTACATGAGTCTCTGATCTGATCTGATCTGATCTGATCTG 642
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 GTCTCCACACAGCTTACATGAGTCTCTGATCTGATCTGATCTGATCTGATCTG 659
QY 643 TTACTGTGCAAGAGGTTACGAGTGTGATCTGATCTGATCTGATCTGATCTGATCT 702
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 TTCTCTGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 719
QY 703 CACC 706
    ||| |||
Db 720 CACC 723

RESULT 10
073679
ID 073679 standard: DNA; 729 BP.
AC 073679;
DC 09-MAY-1995 (first entry)
DE Fv(GP-4) immunosuppressive.
KW MAb: monoclonal antibody; hybridoma: Interleukin-2; IL-2;
    Fv: antibody variable region; GP-4; Fv(GP-4);
    immunosuppressive; ss.
OS Mus sp.
PN EP-621338-A.
PD 26-OCT-1994.
PE 21-APR-1994; 106257.
PR 21-APR-1993; JP-094491.
PR 07-MAR-1994; JP-036065.
PA (AJIN) AJINOMOTO KK.
PI Hamura J, Kanayama Y, Nakazawa H, Shimamura T, Sugamura K;
    Takekoshi T;
    WPI: 94-325948/41.
DR P-SDB: R60781.
PT Immunosuppressant polypeptide - has ability to block
    Interleukin-2 response
```

```
PS Claim 18: Page 29; 37pp; English.
CC MAb capable of binding to the gamma chain of the IL-2 receptor, and
CC thus of blocking the IL-2 response, is produced by mouse hybridoma
CC line GP-4 (FERM BP-4640). DNA encoding the variable region of
CC this MAb was expressed in E. coli, yielding Fv(GP-4) with
CC immunosuppressive activity.
SQ Sequence 729 BP; 199 A; 169 C; 182 G; 179 T;

Query Match
48.9%; Score 348.4; DB 1; Length 729;
Best Local Similarity 71.5%; Pred. No. 8e-84;
Matches 519; Conservative 0; Mismatches 191; Indels 16; Gaps 4;

QY 1 ATGACCTGAGCTGACCCAGCTCTCCAGATCTGTCTGATCTCCAGGGGAAAGTA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGATATCTGCTGACACAGCTCTCCAGCTCTATCTGATCTGAGGAAAGTGT- 60
QY 61 CACAATGACTTCAGAGGCCACCCCAAGTGA---AGTTACATGCACTGTATACAGAA 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 CACCATCTCTGAGAGCCAGCTGGAATATTCATTAATTAATTAATTAATTAATTAAT 119
QY 118 GCCAGATCTCTCCCAACCTTGATTTATACACATCCAACTGCTCTGAGTCCC 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 ACAGGAAATCTCTGATGCTCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 179
QY 178 TGTGCTGCTGAGTGGGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 ATCAAGTGTGAGTGGAGTGGATGAGAACACATATCTCTCAAGTCAACAGCCGTGCA 239
QY 238 GCGTGAAGATGCTGCCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 GCGTGAAGATTTGGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 299
QY 298 AGGGGGGTCCAGCTGGAATTAAGGTCTA---CCTGCTGTGCTGCTGCTGCTGCTGA 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 TGGAGGAGCAAGCTGAGCTCAAGTCAAGATCTCAAGATCTGCTGCCAATCAAA 359
QY 355 AGGTAAAGTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 AAGAGCAGAGTCAAAATGAGAGTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 419
QY 415 GAAGATATCTGCAAGCTTGTGTTACTGATCTGATCTGATCTGATCTGATCTGATCTG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 GAAGATATCTGCAAGCTTGTGTTACTGATCTGATCTGATCTGATCTGATCTGATCTG 479
QY 475 GCAGAGCATGGAAGAAAGCTTGATGATGATGATGATGATGATGATGATGATGATGAT 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GCAGAGCATGGAAGAAAGCTTGATGATGATGATGATGATGATGATGATGATGATGAT 539
QY 535 TAACTCAACAGAGTTCAGAGGCAAGCCACATTAATTAATTAATTAATTAATTAATTA 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 540 TAGCTCAACAGAGTTCAGAGGCAAGCCACATTAATTAATTAATTAATTAATTAATTA 599
QY 595 AGCCTACATGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 600 AGCCTACATGAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 659
QY 655 GAGG-----CTTACGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 660 AGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 719
QY 706 GATCTC 711
    ||| |||
Db 720 TGTCTC 725

RESULT 11
041069
ID 041069 standard: DNA; 848 BP.
AC 041069;
DC 25-AUG-1993 (first entry)
DE 26-10 svf coding sequence.
KW Heavy; light; variable; VH; VL; region; antidiigitoxin; monoclonal;
```









```

Db 541 GACCATGCAATTCACGTGGGTGAACAGAACCCCTGACAGGGGCTGGAATGATTTGATAT 600
QY 512 ATTAATCCTTACATGGTGTACTACTACACAGAAAGTTCAAGGGCAAGCCACATTT 571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 TTTCTCCCCGAAATGATGATTTTAAATACAAATGAGAGGTTCAAGGGCAAGCCACACTG 660
QY 572 ACTGTAGACAAGTGTGACGACAGCCCTACATGGAGCTCCTCAGTGTGACATCTGAGGAC 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 ACTGACAGCAAAATCTCCAGCACTGCTCCTACGTGAGCTCAACAGCCTGACATCTGAGGAT 720
QY 632 TCTGAGTCTATTACTGTGCAAG 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 TCTGAGTCTATTCTGTACAG 743

```

Search completed: May 13, 1999, 09:55:24  
 Job time: 1182 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 09:54:13 ; Search time 736.46 Seconds

(without alignments)  
3497.589 Million cell updates/sec

Title: US-08-704-178-2

Perfect score: 1 ATGCAGCTGACCCAGTCTCC.....CGGTACCGTCTCTCTATAA 720

Sequence:

Scoring table: IDENTITY\_NUC

Searched: 808301 seqs, 1788773984 residues

Database :

GenEmbl.\*  
1: gb\_bal:\*  
2: gb\_bal2:\*  
3: gb\_in:\*  
4: gb\_in:\*  
5: gb\_ov:\*  
6: gb\_ov:\*  
7: gb\_pat:\*  
8: gb\_pat:\*  
9: gb\_pat2:\*  
10: gb\_pat2:\*  
11: gb\_pat2:\*  
12: gb\_pat2:\*  
13: gb\_pat2:\*  
14: gb\_pat2:\*  
15: gb\_pat2:\*  
16: gb\_pat2:\*  
17: gb\_pat2:\*  
18: gb\_pat2:\*  
19: gb\_pat2:\*  
20: gb\_pat2:\*  
21: gb\_pat2:\*  
22: gb\_pat2:\*  
23: gb\_pat2:\*  
24: gb\_pat2:\*  
25: gb\_pat2:\*  
26: gb\_pat2:\*  
27: gb\_pat2:\*  
28: gb\_pat2:\*  
29: gb\_pat2:\*  
30: gb\_pat2:\*  
31: gb\_pat2:\*  
32: gb\_pat2:\*  
33: gb\_pat2:\*  
34: gb\_pat2:\*  
35: gb\_pat2:\*  
36: gb\_pat2:\*  
37: gb\_pat2:\*  
38: gb\_pat2:\*  
39: gb\_pat2:\*  
40: gb\_pat2:\*  
41: gb\_pat2:\*  
42: gb\_pat2:\*  
43: gb\_pat2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Query Match Length DB ID Description

1	720	100.0	720	6	132407	132407 Sequence 2
2	433.4	60.2	711	6	132406	132406 Sequence 1
3	433.4	60.2	711	15	S39590	S39590 anti-erbB2
4	432	60.0	732	6	I13035	I13035 Sequence 1
5	432	60.0	732	27	E10361	E10361 cDNA encod1
6	404.4	56.2	738	15	SYNMR31GA	M34000 Synthetic m
7	383.8	53.3	720	6	E13589	E13589 DNA encodin
8	333.8	46.4	828	13	MM295478	295478 Mus musculu
9	327.6	45.5	822	13	MM295477	295477 Mus musculu
10	326.4	45.3	735	6	E13598	E13598 DNA encodin
11	321.2	44.6	741	6	I08678	I08678 Sequence 10
12	311.6	43.3	753	6	I08679	I08679 Sequence 12
13	299	41.5	729	6	I1036	I1036 Sequence 3
14	299	41.5	729	27	E10362	E10362 cDNA encod1
15	297.4	41.3	744	15	SYNDA44X	L17037 Synthetic s
16	293.6	40.8	803	6	AR012824	AR012824 Sequence
17	292.6	40.6	797	6	AR012822	AR012822 Sequence
18	289	40.1	819	6	A18692	A18692 Synthetic n
19	288.8	39.8	861	13	MCCEAHC	X52769 M.musculus
20	286.8	39.8	711	6	I08677	I08677 Sequence 8
21	283.4	39.4	401	13	MUSIGLH	M64623 Mouse reart
22	282.8	39.3	444	13	MUSIGHD2	M13328 Mouse Ig mu
23	282.8	39.3	444	13	MUSIGHRA	M13329 Mouse Ig mu
24	282.4	39.2	417	13	MUSIGHNP	M19570 Mouse Ig ac
25	280.8	39.0	511	13	MUSIGHBC	K00608 Mouse Ig ac
26	280.2	38.9	318	13	AF057545	AF057545 Mus muscu
27	278	38.6	321	13	MUSIGLAV	M20464 Mouse (clon
28	277.4	38.5	321	13	MUSIGLAV	M20465 Mouse (clon
29	273.4	38.0	299	13	MM055616	U55616 Mus musculu
30	273.2	37.9	321	13	MUSIGLAX	M20467 Mouse (clon
31	272.6	37.9	321	13	AF021870	AF021870 Mus muscu
32	271.8	37.8	723	13	AF035617	AF035617 Mus muscu
33	270.4	37.6	369	13	MMVRHC	X93024 M.musculus
34	270	37.5	501	13	MUSIGHBH	K00687 Mouse Ig ac
35	270	37.5	501	13	MMIGVK1	X00894 Mouse mRNA
36	269.6	37.4	390	13	MUSIGHB	M13832 Mouse Ig ac
37	269.6	37.4	304	13	MM055603	U55603 Mus musculu
38	269.6	37.4	315	13	MMIG1KL2	X02561 Mouse mRNA
39	269.6	37.4	420	13	MMIGNP2	X02563 Mouse mRNA
40	269.6	37.4	354	13	MDIGMAP	222078 M.domesticu
41	269.2	37.4	321	13	MUSIGLAV	M20466 Mouse (clon
42	269.2	37.4	456	13	MUSIGHVAC	X14624 Murine MRL-
43	268.4	37.3	351	13	MUSIGHNP	L22322 Mouse reart
44	268.4	37.3	414	13	MUSIGHNP	M12749 Mouse Ig ac
45	268.4	37.3	414	13	MMIGNP7	X02568 Mouse mRNA

ALIGNMENTS

RESULT 1	132407	720 bp	DNA	PAT	07-JAN-1997
LOCUS	132407	Sequence 2 from patent US 5587458.			
DEFINITION	132407				
ACCESSION	g1823198				
NID					
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 720)				
AUTHORS	King/C.Richter, Kasprzyk, P.G. and Bird, R.E.				
TITLE	Anti-erbB-2 antibodies: combinations thereof, and therapeutic and				
JOURNAL	Patent: US 5587458-A 2 24-DEC-1996;				
FEATURES	Location/Qualifiers				
SOURCE	1..720				
BASE COUNT	184 a 178 c 179 g 179 t				
ORIGIN	/organism="unknown"				

Query Match 100.0%; Score 720; DB 6; Length 720;

Best Local Similarity 100.0%; Pred. No. 1.3e-211;  
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACGCTGACCCAGCTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAAGGTCAACATG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 ATGACGCTGACCCAGCTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAAGGTCAACATG 60
QY 61 ACCGTCAGTGCAGCTCAAGTAAATGACATGCACTGGTATCAGCAGAAAGTCAAGCACC 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ACCGTCAGTGCAGCTCAAGTAAATGACATGCACTGGTATCAGCAGAAAGTCAAGCACC 120
QY 121 TCCCCCAAACTGGGTTATATGACATCAACATGCGCTTCGAGATGCCAGAGTGCCTTC 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 TCCCCCAAACTGGGTTATATGACATCAACATGCGCTTCGAGATGCCAGAGTGCCTTC 180
QY 181 AGTGGCAGTGGGTCTGGAAACTCTTACTCTCAGCATCAGCAGTGGAGCTGAAAT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 AGTGGCAGTGGGTCTGGAAACTCTTACTCTCAGCATCAGCAGTGGAGCTGAAAT 240
QY 241 GCTGCCATTATTTATTTATATCAGGGAGTGGTACCACTTCAGCTTCGGCTGGGGACA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 GCTGCCATTATTTATTTATATCAGGGAGTGGTACCACTTCAGCTTCGGCTGGGGACA 300
QY 301 AAGTTGGAAATAAAGTCTTACCTCCGATCTGTAATCTTCTGAAGTAAAGTGTG 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 AAGTTGGAAATAAAGTCTTACCTCCGATCTGTAATCTTCTGAAGTAAAGTGTG 360
QY 361 CAGCTGCACAGCTCTGGGGTTGAGCTTTCGAGAGGGGCGCTTGAAGTGTGCTGC 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 CAGCTGCACAGCTCTGGGGTTGAGCTTTCGAGAGGGGCGCTTGAAGTGTGCTGC 420
QY 421 AAAGCTTCTGACTTCAACATTAAGACATATTAATCCACTGGGTGAAGCAGAGGCGCTGA 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 AAAGCTTCTGACTTCAACATTAAGACATATTAATCCACTGGGTGAAGCAGAGGCGCTGA 480
QY 481 CAGGGCGCTGGAATGATTTGGATGGATTCCTGAGAAATGTAATCTTATATGACCG 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 481 CAGGGCGCTGGAATGATTTGGATGGATTCCTGAGAAATGTAATCTTATATGACCG 540
QY 541 AAATTCAGAGGCAAGGCGCAGTAAACAGACAGACATCTCCAGCGGCTACCTTCAG 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 541 AAATTCAGAGGCAAGGCGCAGTAAACAGACAGACATCTCCAGCGGCTACCTTCAG 600
QY 601 CTCAGACGCTGACATCTGAGACACTGCGCTATTTACTGTCTTATTTACTACTAT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 601 CTCAGACGCTGACATCTGAGACACTGCGCTATTTACTGTCTTATTTACTACTAT 660
QY 661 AGTGGTACTATGCTATGTAATCTGAGGCTCAAGGAACTCGGCTACCGCTCTCTCATTA 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 661 AGTGGTACTATGCTATGTAATCTGAGGCTCAAGGAACTCGGCTACCGCTCTCTCATTA 720
```

RESULT 2  
LOCUS 132406 711 bp DNA PAT 07-JAN-1997  
DEFINITION Sequence 1 from patent US 5587458.  
ACCESSION 132406  
NID 91823197  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 711)  
AUTHORS King, C. Richter, Kasprzyk, P. G. and Bird, R. E.  
TITLE Anti-erbB-2 antibodies, combinations thereof, and therapeutic and  
diagnostic uses thereof  
JOURNAL Patent: US 5587458-A 1 24-DEC-1996;  
FEATURES Location/Qualifiers  
source 1..711  
BASE COUNT 175 a 182 c 190 g 164 t  
ORIGIN -

Query Match 60.2%; Score 433.4; DB 6; Length 711;  
Best Local Similarity 79.6%; Pred. No. 3.4e-123;  
Matches 512; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

```
QY 2 TGCACGTACCCAGCTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAAGGTCAACATG 61
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8 TGCACGTACCCAGCTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAAGGTCAACATG 67
QY 62 CTGAGTGCAGCTCAAGTAAATGACATGCACTGGTATCAGCAGAAAGTCAAGCACC 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 62 CTGAGTGCAGCTCAAGTAAATGACATGCACTGGTATCAGCAGAAAGTCAAGCACC 121
QY 68 CTGAGGAGGCCACCCCAAGTAAATGACATGCACTGGTATCAGCAGAAAGTCAAGCACC 127
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 68 CTGAGGAGGCCACCCCAAGTAAATGACATGCACTGGTATCAGCAGAAAGTCAAGCACC 127
QY 122 CCCCCAACTGGGTTATATGACATCAACATGCGCTTCGAGATGCCAGAGTGCCTTC 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 122 CCCCCAACTGGGTTATATGACATCAACATGCGCTTCGAGATGCCAGAGTGCCTTC 181
QY 128 CCCCCAACTGGGTTATATGACATCAACATGCGCTTCGAGATGCCAGAGTGCCTTC 187
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 128 CCCCCAACTGGGTTATATGACATCAACATGCGCTTCGAGATGCCAGAGTGCCTTC 187
QY 182 GTGGCAGTGGGTCTGGAAACTCTTACTCTCAGCATCAGCAGTGGAGCTGAAATG 241
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 182 GTGGCAGTGGGTCTGGAAACTCTTACTCTCAGCATCAGCAGTGGAGCTGAAATG 241
QY 242 CTGCCACTTATTTATTTATATCAGGGAGTGGTACCACTTCAGCTTCGGCTGGGGACA 301
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 242 CTGCCACTTATTTATTTATATCAGGGAGTGGTACCACTTCAGCTTCGGCTGGGGACA 301
QY 248 CTGCCACTTATTTATTTATATCAGGGAGTGGTACCACTTCAGCTTCGGCTGGGGACA 307
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 248 CTGCCACTTATTTATTTATATCAGGGAGTGGTACCACTTCAGCTTCGGCTGGGGACA 307
QY 302 AAGTTGGAAATAAAGTCTTACCTCCGATCTGTAATCTTCTGAAGTAAAGTGTG 361
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 302 AAGTTGGAAATAAAGTCTTACCTCCGATCTGTAATCTTCTGAAGTAAAGTGTG 361
QY 362 AGCTGCACAGCTCTGGGGTTGAGCTTTCGAGAGGGGCGCTTGAAGTGTGCTGC 421
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 362 AGCTGCACAGCTCTGGGGTTGAGCTTTCGAGAGGGGCGCTTGAAGTGTGCTGC 421
QY 368 AGCTGCACAGCTCTGGGGTTGAGCTTTCGAGAGGGGCGCTTGAAGTGTGCTGC 427
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 368 AGCTGCACAGCTCTGGGGTTGAGCTTTCGAGAGGGGCGCTTGAAGTGTGCTGC 427
QY 422 AAGCTTCTGACTTCAACATTAAGACATATTAATCCACTGGGTGAAGCAGAGGCGCTGAC 481
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 422 AAGCTTCTGACTTCAACATTAAGACATATTAATCCACTGGGTGAAGCAGAGGCGCTGAC 481
QY 482 AAGGCGCTGGAATGATTTGGATGGATTCCTGAGAAATGTAATCTTATATGACCG 541
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 482 AAGGCGCTGGAATGATTTGGATGGATTCCTGAGAAATGTAATCTTATATGACCG 541
QY 488 AGAACCCTGAGTGGATTTGGATGGATTCCTGAGAAATGTAATCTTATATGACCG 547
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 488 AGAACCCTGAGTGGATTTGGATGGATTCCTGAGAAATGTAATCTTATATGACCG 547
QY 542 AATTCAGAGGCAAGGCGCAGTAAACAGACAGACATCTCCAGCGGCTACCTTCAG 601
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 542 AATTCAGAGGCAAGGCGCAGTAAACAGACAGACATCTCCAGCGGCTACCTTCAG 601
QY 548 AGTTCAAGGGCAAGGCGCAGTAAACAGACAGACATCTCCAGCGGCTACCTTCAG 607
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 548 AGTTCAAGGGCAAGGCGCAGTAAACAGACAGACATCTCCAGCGGCTACCTTCAG 607
QY 602 TCAGCAGCTGACATCTGAGACACTGCGCTATTTACTGTGC 644
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 602 TCAGCAGCTGACATCTGAGACACTGCGCTATTTACTGTGC 644
QY 608 TCCTCAGTCTGACATCTGAGACACTGCGCTATTTACTGTGC 650
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 608 TCCTCAGTCTGACATCTGAGACACTGCGCTATTTACTGTGC 650
```

RESULT 3  
LOCUS S39590 711 bp SYN 10-FEB-1993  
DEFINITION anti-erbB2 immunotoxin antigen binding region [mice, Other  
Synthetic Recombinant Partial, 711 nt].  
ACCESSION S39590  
NID 9251113  
KEYWORDS  
SOURCE Mus sp.  
ORGANISM Mus sp.  
REFERENCE 1 (bases 1 to 711)  
AUTHORS Batra, J. K., Kasprzyk, P. G., Bird, R. E., Pastan, I. and King, C. R.  
TITLE Recombinant anti-erbB2 immunotoxins containing pseudomonas exotoxin  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (13), 5867-5871 (1992)  
MEDLINE 92335198  
REMARK GenBank staff at the National Library of Medicine created this  
entry [NCBI g108547] from the original journal article.  
FEATURES This sequence comes from Fig. 1.  
source 1..711  
Location/Qualifiers



	RESULT	5	
Db	EI0361	standard; RNA; ROD; 732 BP.	732 TCCTCA
XX	EI0361		
AC	EI0361;		
XX	d1108698		
XX	08-OCT-1997 (Rel. 52, Created)		
DT	08-OCT-1997 (Rel. 52, Last updated, Version 1)		
XX	cDNA encoding a monoclonal antibody against human Interleukin 2		
DE	receptor gamma chain.		
XX	JP 1995313188-A/1.		
KW	Mus sp.		
XX	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia;		
OC	Sclurognathii; Muridae; Murinae; Mus.		
XX	[1]		
RN	1-732		
RP	Shimamura T., Hamuro J., Nakazawa H., Kanayama Y., Sugamura K.,		
RA	Takeshita T.;		
RT	"IMMUNOSUPPRESSIVE AGENT";		
RL	Patent number JP 1995313188-A/1, 05-DEC-1995.		
XL	AJINOMOTO CO INC., SUGAMURA KAZUO.,		
XX			
OS	Mus sp. (mouse)		
CC	PN JP 1995313188-A/1		
CC	PD 05-DEC-1995		
CC	PF 21-APR-1994 JP 1994082836		
CC	PR 21-APR-1993 JP 93P 94491, 07-MAR-1994 JP 94P 36065		
CC	PI SHIMAMURA TOSHIAKI, HAMURO JUNJI, NAKAZAWA HARUMI,		
CC	PI KANAYAMA YUKA,		
CC	PI SUGAMURA KAZUO, TAKESHITA TOSHITACHI		
CC	C12P21/08,A61K39/395,A61K39/395,C12N1/21,C12N5/20,		
CC	C12N15/13//C12N15/06,		
CC	PC (C12P21/08,C12P1:19),(C12P21/08,C12P1:91),(C12N1/21,C12P1:19);		
CC	CC CC strandness: Double;		
CC	CC topology: Linear;		
CC	FH Key location/Qualifiers		
CC	FH source 1..732		
CC	FT /organism="Mus sp."		
CC	FT /cell_type="hybridoma"		
CC	FT /cell_line="Gp-2"		
CC	FT mat_peptide 1..732		
CC	/product="anti-IL-2 receptor gamma chain"		
XX			
TH	Key Location/Qualifiers		
FT	source 1..732		
FT	/organism="Mus sp."		
SO	Sequence 732 BP; 180 A; 204 C; 182 G; 166 T; 0 other;		
<hr/>			
	Query Match	60.0%; Score 432; DB 27; Length 732;	
	Best Local Similarity	76.7%; Pred. No. 9.3e-123;	
	Matches 557; Conservative 0; Mismatches 160; Indels 9; Gaps 2;		
OY	1 ATGCAGTGCACCAGTCATGCCAAATCATGTGTGCATCTCCAGGGGAAAAGTCAACCATG 60		
Dd	7 ATCCGTGTGACCCAGTCATCCATCAATCATGTGTGCATCTCTAGGGGAACGGGTCAACATG 66		
	ACCTGCACTGGCAGCGTCAAGTGTA-----ATAACATCCACTGGTATACGAGAAGTCA 114		
Db	67 ACCTGCACTGGCAGCGTCAAGTGTAAGTTCACTTCACTTCACTGGTACCAAGAGCCA 126		

QY	115	AGCAACCTCCCAACACTGGTGTATTGACACATCCAAACCTGGCTTCGGAGTCCCAAGT	174
Db	127	GGATCTCCCAACACTGGATTTATAGCAATCCAACTGGCTTCGGAGTCCCAAGT	186
QY	175	CGCTTCAGTGGCACTGGGTCTGGAACACTCTACCTCTCCAGATCAGCAGATGGAGGCT	234
Db	187	CGCTTCAGTGGCACTGGGTCTGGAGCCTTACCTCTCCAAATCAGCAGATGGAGGCT	246
QY	235	GAAGATGCTGCCACTTATTATTGTTATCAGGGAGTGGGTACCCATTACGTTGGGCTCG	294
Db	247	GAAGATGCTGCCACTTATTATTGCTCCACACAGTATCATCTTCCCGCTCACGTTGGTGCT	306
QY	295	GGGACAAAGTGGAAATTAAGGTTCTA---CCGCCGATCTGGAAATCTTCGAAAGCT	351
Db	307	GGGACAAAGTGGAGCTCAAAAGTCAGCAAAATCTCAGGATCTGGCTCCGAATCCAAAGC	366
QY	352	AAAGGTGTGACGCTGCAGCAGCTCGGGGTGAGCTTGTCCGAGGAGGGGCGCTTAGTCAAG	411
Db	367	ACGAGGTCAAACTCGAGAGACTCTGGAATCTGAGTGGAGGCTGGAGCTTCAGTGAAG	426
QY	412	TTGTCTCGCAAAAGCTTCTGACTTCAACATTAAGAATAATTATTCACCTGGGTGAAGCAG	471
Db	427	CTGTCTCGCAAGGCTTCTGGCTACACATTCACCAAGCTACTGGATGCACTGGGTGAAGCAG	486
QY	472	AGGCTGGAACAGGGCGCTGGATGGATTGGATGGATTCATCTCTGAGAAATGATTAATCTGT	531
Db	487	AGGATGGACAAAGGCTTGAATGGATGGAAATTTATATCTGTGATGAGTAACTAAC	546
QY	532	TATGACCCGAATTCAGAGGGCAGGCCAGTATTAACAGACAGACATCTCCACGGGCGC	591
Db	547	TACATGATGAAGTTCAGAGGACAGGCCACACTGATCTAGACACATCTCCACGACAGCC	606
QY	592	TACCTTCAGCTCAGCAGGCTTGACATCTGAGGACACACTGCGCTATTACTGTCTTCTTAT	651
Db	607	TACATGACACTCAGCAGGCTGACATCTGAGGACACTGCGCTATTACTGACAAAGAGC	666
QY	652	TACACTATATGCTTACTATATGCTATGCTATGCTACACTGCGGCTCAAGAACTCGCTCACGCTC	711
Db	667	AGCCGGAACCTGGTCTACTATGCTATGCTACACTGCGGCTCAAGAACTCGCTCACGCTC	726
QY	712	TGCTCA 717	
Db	727	TGCTCA 732	
RESULT	6		
SYNAR3IGA			
LOCUS	SYNAR3IGA	738 bp	DNA SYN 15-SEP-1990
DEFINITION	Synthetic mouse Ig heavy and light chain antibody (OVBJ), partial cds.		
ACCESSION	M34000		
NID	g207974		
KEYWORDS	C-region.		
SOURCE	Mouse hybridoma, cDNA to mRNA, and E.coli vector DNA.		
ORGANISM	artificial sequence.		
REFERENCE	1 (bases 1 to 738)		
AUTHORS	Chaudhary,V.K., Batra,J.K., Gallo,M.G., Willingham,M.C., Fitzgerald,D.J. and Pastan,I.		
TITLE	A rapid method of cloning functional variable-region antibody genes in Escherichia coli as single-chain immunotoxins		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 87, 1066-1070 (1990)		
MEDLINE	90138938		
FEATURES			
source	Location/Qualifiers		
	1..738		
	/organism="artificial sequence"		
	/db_xref="taxon:29278"		
	<1..>738		
	/note="OVBJ Ig light and heavy chains"		
	/codon_start=1		
	/transl_table=1		
	/db_xref="PID:g207975"		
	/translation="HMEENVLTQSPDAIMASPGEKVTMTGSAISTVSIHMYIQKSSSTS"		







CDS  
/db\_xref="taxon:10090"  
1..822

/codon\_start=1  
/product="IgG1/kappa antibody"  
/db\_xref="PID:e1167935"  
/db\_xref="PID:92108311"  
/translation="MDVMTOTPLSPVSLDQASISCRSSQTIHSHNNTYLEFLO  
KPGSPKLLIKVSNRESGVPDRESGSDTFLKISRVEHEDLGVTFCQSHVFP  
TFGGGTLEIRGGSGGGSGGSEVQLOQSGALVRSASVLSCTASDFNI  
KDYIHWKORPEQLEWIGMDPKNGSDTDAKPKQKATMTADTSNTAYLOLSLT  
SEDAVYCCSAVDEPDVARGGQTLTVSSATTPGAAMEQKLSIEDINDKDEL"

misc\_feature

4..339  
/note="light chain variable domain (VL) coding sequence"

misc\_feature

340..393  
/note="Linker peptide coding sequence"

misc\_feature

394..741  
/note="Heavy chain variable domain (VH) coding sequence"

misc\_feature

742..756  
/note="heavy chain first constant domain (CH1) fragment  
coding sequence"

misc\_feature

769..801  
/note="c-myc tag coding sequence"

misc\_feature

808..819  
/note="KDEL coding sequence"

misc\_feature

214 a. 203 c 221 g 184 t

BASE COUNT  
ORIGIN

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.3%; Score 326.4; DB 6; Length 735;  
Best Local Similarity 67.6%; Pred. No. 3.6e-90;

Matches 491; Conservative 0; Mismatches 226; Indels 9; Gaps 2;

RESULT 10  
E13598

LOCUS E13598 735 bp DNA PAT 27-APR-1998

DEFINITION DNA encoding a single chain Fv antibody of anti-T3 antibody TTL.

ACCESSION E13598

KEYWORDS JP 1997220092-A/1.

SOURCE JP 1997220092-A/1.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 735)

AUTHORS Eklida, T., Yasukawa, K., Imanaka, T. and Takagi, M.

TITLE PRODUCTION OF SINGLE-STRAND FV ANTIBODY

JOURNAL Patent: JP 1997220092-A 1 26-AUG-1997;

COMMENT TOSOH CORP

OS None

OC Artificial sequences.

PN JP 1997220092-A/1

PD 26-AUG-1997

PF 15-FEB-1996 JP 1996027622

PI EKIDA TEIJI, YASUKAWA KIYOSHI, IMANAKA TADAYUKI, PI

MASAHIRO

PC C12N1/09, C12N1/21, C12P2/08, (C12N1/21, C12R1/19), (C12P2/08,

PC C12R1/19);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

FE Key

FT source

FT 1..735

FT location/Qualifiers

FT /organism='Artificial sequences'.

FEATURES

source

1..735

/organism='unclassified'

/db\_xref="taxon:32644"

BASE COUNT

ORIGIN

Query Match

45.3%; Score 326.4; DB 6; Length 735;  
Best Local Similarity 67.6%; Pred. No. 3.6e-90;

Matches 491; Conservative 0; Mismatches 226; Indels 9; Gaps 2;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;

Query Match

45.5%; Score 327.6; DB 13; Length 822;  
Best Local Similarity 73.9%; Pred. No. 1.5e-90;

Matches 436; Conservative 0; Mismatches 139; Indels 15; Gaps 1;



QY	300	AAAGTTGGAAATAAAGTTTACCTACCGGAGTGGTAAATCTTTGAAGCTAAAGTGT	359
Db	312	CAAGCTGGAGCGAAGAGGACTGTGCTCCAGCAGATAAAACTTAACGTTAAAGCAGATG	371
QY	360	G-----CACTGCAGCAGCTGTGGGGTTGAGCTTGTCCGAGAGGGGGCTT	404
Db	372	GGGCAAGATTAGTACCTACCTGGTGGAGTCTGGGGGAGACTTAGTGAAGCTGGAGGTC	431
QY	405	AGTCAAGTTTGCTGCCAAAGCTTCTGACTTTCACATTTAAAGACTATTATATCCATGGGT	464
Db	432	CCTGAAGATCTCTGTGCGAGCATCTGGATTCTACTTTCATTAGCTATGAGCATGCTTGGGT	491
QY	465	GAAGCAGAGCGCTTGAAACAGGCGCTGGAAATGATGTGATGATTCATCTCAGATGTGA	524
Db	492	TGCGCAGACTCCAGACAAAGAGGCTGGAGTGGGTCCGCAACNNAGNAGTGTGTACTTA	551
QY	525	TACTGTATATGACCCGAAATTTCCAGGGCAAGGCCACTATATACAGCAGACATCTCCAA	584
Db	552	CACCTACCTATCCAGACAGTGAAGGGGGGATTCACCACTCTCCAGAGACATGCCAAGAA	611
QY	585	CGCGGCTACCTTCAGCTCAGCAGCCCTGACATCTGAGGACACACGCCGTCTATTACTGTGC	644
Db	612	CACCTGTACCTGCAAAATAGCGGGTGTGAAGTGTGAGGACACAGCCATGTATTACTGTGC	671
QY	645	--TCTTTATTACTACTATAGT-----GCTTACTATGCTATGTACTACTGSGGTCAAGG	695
Db	672	AAGAGCGATTTACTAGGTTGACTTACGAGATTACTATGTATGTAGTACTACTGSGGTCAAGG	731
QY	696	AACCTGGGTCACCGTCTCTCTCA	717
Db	732	AACCTCAGTCACCGTCTCTCTAA	753
RESULT	13		
LOCUS	131036		
DEFINITION	Sequence 3 from patent US 5582826.		
ACCESSION	131036		
NID	91821827		
KEYWORDS			
SOURCE			
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 729)		
AUTHORS	Shimamura,T., Hamuro,J., Nakazawa,H., Kanayama,Y., Sugamura,K. and Takeshita,T.		
TITLE	Monoclonal antibodies which bind the gamma chain of human interleukin-2 receptor		
JOURNAL	Patent: US 5582826-A 3 10-DEC-1996;		
FEATURES	1..729		
BASE COUNT	199 a 169 c 182 g 179 t		
ORIGIN	/organism="unknown"		

[illegible]

OY	178	TTTACGTGCATGGGCTGTGAAATCTTACTCTCTACAACTACAGACATGAGGGCTGAA	237
Db	187	TTTCAGTGGCACTGGATATAGGAACCAATATTTCTCTACATGATCAACCCCTGCACCTGAA	246
OY	238	GATGCTGCACCTATTTATTTGTTATCAGGGAGTGGGTACCATTCACTGGCTGGGG	297
Db	247	GATTTTGGGAGTTATTTACTGTCCAAATTTTGGAGTACTCGTGGAGCTTGGTGGAGGG	306
OY	298	ACAAAGTTGGAAATAAAAAGTTCTA--CCTCCGGATCTGGTAAATCTTGTGAAGTAAA	354
Db	307	ACCAAGCTGGAGCTCAAAAGTCAGAAATCCTCAAGATCTGGCTCCGATCCAAAGACAG	366
OY	355	GGTGGGAGCTGCAGCAGTCGTGGGGCTTGACCTGTCCGAGAGGGGGCTTACGCAAGTTG	414
Db	367	CAGGTCAAACTCGAGAGATCTGGACCTGCAGCTGGTGAAGCTGGGGCTTGTAGTGAAGATA	426
OY	415	TTCTGCAAAAGCTTTGTGACTTCAACATTAAAGACTATATATATCCACTGGGTGAAGCAGAG	474
Db	427	TTCTGCAAAAGCTTTGTGTTACTATTCATCGCTACTACATGCACTGGGTGAAGCAAAAGC	486
OY	475	CCTGACACAGGGCCTTGGATGGATTGATGATTCATCTCGAAGATGGTAAATCTGTATAT	534
Db	487	CATGTAAGAGCCTTGTGATGGATTTGGACGTATTAATCTTACAAATGGTCTACTAGCTAC	546
OY	535	GACCGGAAATTCGAGGGCAAGGGCCAGATATAACGACACACATCTCCACAGGGGGCTAC	594
Db	547	AACCAAGATTTTCAAGGACAAAGGGCCAGCTTGACTGTATATAGTCTCTCCACACAGCCTAC	606
OY	595	CTTCAGCTCAGACACCTGACATCTGAGAGCACTGCCCTCTATTACTGTGCTTCTATTAC	654
Db	607	ATGGAGCTCCACACCCGTGACATCTGAGAGACTCTGCAGTCTATTATCTGTGCAAGAGAAAT	666
OY	655	TACTATAGTGCTTACTATAGCTATGCTATGCTACTACTGGGGTTCAGAGAACTGGTACCGTCC	714
Db	667	TACTACGCTAGTACTACAGGGTGTGCTTACTACGGGGCCAAAGGACTGTGGTACACTGTCTCT	726
OY	715	TCA 717	
Db	727	GCA 729	

RESULT	14		
ID	E10362	standard; RNA; ROD; 729 BP.	
XX	E10362		
XX	E10362;		
XX			
NI	d1108699		
XX			
DT	08-OCT-1997 (Rel. 52, Created)		
DT	08-OCT-1997 (Rel. 52, Last updated, Version 1)		
XX			
DE	cDNA encoding an monoclonal antibody against human Interleukin-2		
DE	receptor gamma chain.		
XX			
KW	JP 1995313188-A/2.		
XX			
Mus sp.			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia.		
OC	Sciurognathi; Muridae; Murinae; Mus.		
XX			
[1]			
RN	1-729		
RP	Shimamura T., Hamuro J., Nakazawa H., Kanayama Y., Sugamura K.,		
RA	Takeshita T.;		
RT	"IMMUNOSUPPRESSIVE AGENT";		
RL	Patent number JP 1995313188-A/2, 05-DEC-1995.		
XX			
XX	AJINOMOTO CO INC, SUGAMURA KAZUO.,		
OS	Mus sp. (mouse)		
CC	JP 1995313188-A/2		
CC	PD 05-DEC-1995		
CC	PF 21-APR-1994 JP 1994082836		
CC			

CC	PR	21APR-1993 JP 93P	94491.07-MAR-1994 JP 94P	36065
CC	PI	SHIMAMURA TOSHIAKI,	HAMURO JUNJI,	NAKAZAWA HAROMI,
CC	CC	KANAYAMA YUKA,		
CC	PI	SUGAMURA KAZUO,	TAKEISHITA TOSHICHI	
CC	PC	C12P21/08,A6IK39/395,	A6IK39/395,C12N1/21,C12N5/20,	
CC	CC	C12N15/13//C12N15/06,		
CC	PC	(C12P21/08,C12R1:19),	C12P21/08,C12R1:91),(C12N1/21,C12R1:19);	
CC	CC	key	location/Qualifiers	
CC	FH			
CC	FT	source	1. .729	
CC	CC	FT	/organism="Mus sp."	
CC	FT	FT	/cell_type="hybridoma"	
CC	CC	FT	/cell_line="GP-4"	
CC	CC	FT	1. .729	
CC	CC	FT	/product="anti-IL-2 receptor gamma chain"	
XX	XX			
FH	Key	Location/Qualifiers		
FH	source	1. .729		
FT		/organism="Mus sp."		
XX	XX			
SQ	Sequence	729 BP; 199 A; 169 C; 182 G; 179 T; 0 other;		

Query Match	41.5%;	Score 299;	DB 27;	Length 729;
Best Local Similarity	65.3%;	Pred. NO. 1e-81;		
Matches 472;	Conservative	0;	Mismatches 245;	Indels 6;
			Gaps	2;

OY		1	ATGAGCGTACCACAGCTCCAGCAATATGTCTGCATCTCCAGGGGAAAAAGTCACCATG	60
Db		7	ATTTCGCTGACACAGACTCTCCAGCCTCCCTATCTGCATCTGTGGGAGAAACCTGCACCATC	66
OY		61	ACCTGCAGTGCCAGCTCAAGTGTGA--AGTAACATGCAGCTGTTACAGCAAGAACTCAAGC	117
OY		118	ACCTCCCCCAACCTCTGGGTTTTATGACACATCCAAACTGGCTTGTGGAGTCCCAGGTCCG	177
Db		67	AACAATCGACGAAGAATGGGAATATTCACAAATATTTAGCATGTGATTCAGCAAGAAACAGGA	126
OY		127	AAATCTCTCGAAGCTCTGGTTGTATTAATGCAAAAACCTTAGCAGATGGTGTGCCATCAAGG	186
OY		178	TTCAGTGGCAGTGGGTCTGGAACCTCTTACTCTCCAGATCACAGACATGAGAGCTGAA	237
Db		187	TTCAAGTGGCAATGGATGACGAAACAATAATCTCTCAAGATCAACAGACCTGCAGCCTGAA	246
OY		238	GATGCTGCCACTTATATTGTATTACAGGGGAGTGGTACCATTCACAGTTGGCTCGGG	297
Db		247	GATTTGGAGATTTATTACTGTCAACATTTTGGAGTACTCGCTGGAAGTTTGGGAGGAG	306
OY		298	ACAAAGTTGGAAATPAAAGTTCTTA---CCTCCGAGTCTGGTAAATCTTCGAAAGTAAA	354
Db		307	ACCAAAGCTGGAGTCCAAGTGGAGAAATCCTCAGAGATTTGGCTCCGAATCCAAAGACAG	366
OY		335	GGTGTGCAGCTGCACAGCACTGGGGTTGAGCTTGTCCGAGGAGGGGCGTTAGTCAAGTTG	414
Db		367	CAGGTCAAACTCGAGAGTGTGCACTGAGAGCTGGTGAAGCCTGGGGCTTCAGTGAAGATA	426
OY		415	TCCTGCAAAAGTTTGSACTTCAACATTAAACATTAATATCCACTGGGTGAAGAACAGG	474
Db		427	TTCTGCAAGGCTTTCTGGTTACTCAATTCAGCTGGCTACTACTACATGCACCTGGGTGAAGCAAGC	486
OY		475	CCTGAACAGGGCCTGGAATGAGTATGGATGATTCATCCTGAGAAATGTAATACTGTATAT	534
Db		487	CATGTAAGAAGCCTTGAATGATGATGAGAGTATTAATCCTTNCATATGGGTACTAGCTAC	546
OY		535	GACCCGAAATTCAGGGGAAAGGCCAGTATTAACACAGACACATCTCTCCAAAGCGGCTTAC	594
Db		547	AACCAAGATTTTCAAGGAACAAAGCCAGCTGACTAGTAAGTCTCCACAGACAGCCTTAC	606
OY		595	CTTCAAGCTCACAGCTCACATCTGAGAGCACTCCGCTATTAAGTCTGAGCTTTATTATAC	654
Db		607	ATGAGAGCTCACAGCCTGACATCTGAGAGCACTCTGACGTCTTTTACTGTGCAAGAGAAAT	666
OY		655	TACTATAGTCTTACTATGATGTATGTAAGTACTCTGGGGTCAAGGAACCTCGTACCGTCTCC	714

Db	667	TACTACGGTTACTACTACGCGTTGCTTACTTGGGGCCAAAGGACTCTGTCTACTGTCTCT	726
Qy	715	TCA	717
Db	727	GCA	729

RESULT 15

LOCUS	SYNDA4X	744 bp	mRNA	SYN	01-MAR-1996
DEFINITION	Synthetic single-chain Fv fusion protein (DA4.4/212) mRNA, constructed from variable light and heavy chain regions of monoclonal antibody DA4.4, 3' end of cds.				

**KEYWORDS** IgM-binding protein; fusion protein; fusion protein DA4.4/212;

SOURCE Artificial gene cDNA to mRNA.

artificial sequence.

**AUTHORS** Lee, T.K., Rollence, M.L., Halberg, P.L., Oelkuct, M.S., Dodd, S.W.,

**TITLE** Production of engineered IgM-binding single-chain antibodies in *Escherichia coli*

J. Ind. Microbiol. 14 (5), 371-376 (1995)

## FEATURES

### Location/Qualifiers

CDS	<1.	.735	
-----	-----	------	--

```
/codon_start=1
```

```
/function="binds
```

```
/db_xref="PID:g3
```

KPGQSPPELLVFASTH

VNW MKQ R P G R G L E W I G

misc_feature	1. .339
--------------	---------

antibody DA4.4 (

```

300 /note="linker se
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
12
```

```
/note="variable heavy chain derived from mouse monoclonal
```

BASE COUNT	192 a	187 c	186 g	179 t
------------	-------	-------	-------	-------

ORIGIN

Query Match 41.38; Score 297.4; DB 15; Length 744;

Matches 490; Conservative 0; Mismatches 216; Indels 33; Gaps 3;

6 GCTGACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAGGTCACCATGACCTG 65

Db 9 GATGACACAGTCTCCATCCCTGGCTATGTCAGTAGGACAGAGGTCACCTATGAGTTC 68

QY 66 CAGTGCAGCTCAAGTGTAGTACA-----TGCACCTGGTATCA 104

Db 69 CAAGTCCAGTCAGACTCTTTTAATAGTAGCAATCAAAGAAGTATTTGGCCTGTACCA 128

QY 105 GCAGAGTCAAGCACCTCCCCCAACTCTGGTTATGACACATCCAACCTGGCTTCTGG 164

Db 129 GCAGAACCAGGACGCTCTCCTGAACTTCTGGTATACTTGCATCCACTAGGGAATCTGG 188

QY 165 AGTCCAGGTCGCTTCAGTGGCAGTGGCTCTGGAACCTCTTACTCTCTCAGATCAGCAG 224







GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 09:40:43 ; Search time 341.74 Seconds

(without alignments)  
3213.863 Million cell updates/sec

Title: US-08-704-178-2

Perfect score: 720  
Sequence: 1 ATGCAGCTGACCCAGTCTCC.....CGGTACCGTCTCCTCATAA 720

Scoring table: IDENTITY\_NUC

Searched: 2002476 seqs, 762712212 residues

Database :

EST:\*  
1: em\_est1:\*  
2: em\_est2:\*  
3: em\_est3:\*  
4: em\_est4:\*  
5: em\_est5:\*  
6: em\_est6:\*  
7: em\_est7:\*  
8: em\_est8:\*  
9: em\_est9:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: gb\_est6:\*  
16: gb\_est7:\*  
17: gb\_est8:\*  
18: gb\_est9:\*  
19: gb\_est10:\*  
20: gb\_est11:\*  
21: gb\_est12:\*  
22: gb\_est13:\*  
23: gb\_est14:\*  
24: gb\_est15:\*  
25: gb\_est16:\*  
26: gb\_est17:\*  
27: gb\_est18:\*  
28: gb\_est19:\*  
29: gb\_est20:\*  
30: gb\_est21:\*  
31: em\_est10:\*  
32: em\_est11:\*  
33: em\_est12:\*  
34: em\_est13:\*  
35: em\_est14:\*  
36: em\_est15:\*  
37: em\_est16:\*  
38: em\_est17:\*  
39: em\_est18:\*  
40: em\_est19:\*  
41: em\_est19:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	251.2	36.9	396	25	AA691311	vs14f01.r	AA691311 vs14f01.r
2	188.4	26.2	344	22	AA592800	vc25g11.r	AA592800 vc25g11.r

3	149.8	20.8	310	30	A1180569	A1180569 uc70f09.r
4	146.2	20.3	461	31	A1201426	A1201426 gs73c06.x
5	136.6	19.0	387	22	AA569186	AA569186 nm30d10.s
6	133.8	18.6	294	28	A1007196	A1007196 ua73q04.r
7	132.8	17.1	400	20	AA456778	AA456778 zw27f11.r
8	122.2	17.0	393	11	T27593	T27593 EST100653.H
9	119	16.5	328	17	AA295311	AA295311 EST100471
10	118.8	16.5	345	25	AA710291	AA710291 v53a04.r
11	118.6	16.3	403	18	AA345486	AA345486 EST51505
12	117.6	16.3	370	17	AA295093	AA295093 EST100400
13	117.6	16.3	379	21	AA291381	AA291381 EST13847
14	116.6	16.2	379	21	AA291381	AA291381 z744g02.r
15	115.4	16.0	418	21	AA515239	AA515239 dg69c07.s
16	113.8	15.8	363	18	AA367405	AA367405 EST78511
17	113.4	15.8	471	31	A1233978	A1233978 EST230666
18	113	15.7	463	11	R69532	R69532 y182d09.r1
19	112.8	15.7	349	18	AA379044	AA379044 EST91999
20	112.6	15.6	335	17	AA318377	AA318377 EST20820
21	112.4	15.6	398	24	AA423447	AA423447 v680a03.r
22	111.4	15.5	413	17	AA301347	AA301347 EST14279
23	110.6	15.4	345	18	AA335086	AA335086 EST39457
24	109.2	15.2	253	17	AA295941	AA295941 EST101165
25	106.4	14.8	420	30	A1120005	A1120005 uc25e04.r
26	106	14.7	381	11	T29114	T29114 EST69430.Ho
27	105.6	14.7	352	17	AA295786	AA295786 EST100987
28	104.6	14.5	516	11	AA301261	AA301261 EST14181
29	104.6	14.5	516	11	R67559	R67559 y142b11.r1
30	104.6	14.5	346	17	AA300582	AA300582 EST13427
31	104.2	14.5	823	16	AA170256	AA170256 ms87g10.r
32	103.4	14.4	395	11	T29112	T29112 EST69384.Ho
33	102.6	14.2	470	12	H62115	H62115 yu40h01.r1
34	102	14.2	383	11	T29916	T29916 EST99871.Ho
35	101	14.0	405	18	AA360223	AA360223 EST69341
36	97.8	13.6	303	17	AA300891	AA300891 EST14031
37	96.8	13.4	339	25	AA710249	AA710249 v449e01.r
38	96.2	13.4	357	18	AA361497	AA361497 EST71040
39	96	13.3	303	17	AA300788	AA300788 EST13648
40	95.8	13.3	334	17	AA327254	AA327254 EST30547
41	95.4	13.2	210	11	R69482	R69482 y183c03.r1
42	95	13.2	372	25	AA581192	AA581192 nc38b11.r
43	94.8	13.2	432	31	A1268604	A1268604 q047a10.x
44	94	13.1	431	25	AA710970	AA710970 v193h09.r
45	93.2	12.9	332	11	T29342	T29342 EST77181.Ho

## ALIGNMENTS

RESULT 1  
AA691311 396 bp mRNA EST 16-DEC-1997  
DEFINITION vs14f01.r1 Barstead mouse irradiated colon MRLPB7 Mus musculus cDNA  
LOCUS clone 1136201 5' similar to gb:X67211 M.musculus rearranged  
immunoglobulin kappa light chain (MOUSE);, mRNA sequence.

ACCESSION  
AA691311  
NID 92692247

KEYWORDS  
SOURCE  
ORGANISM

house mouse.  
Mus musculus  
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 (bases 1 to 396)  
AUTHORS  
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisler, S., Kucada, T., Lagy, M., Le, M., Martin, V., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
TITLE  
The WashU-HMI Mouse EST Project  
JOURNAL  
Unpublished (1996)  
COMMENT

Contact: Marras M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine



mRNA, complete (mouse);, mRNA sequence.  
 accession A1180569  
 nid 93731207  
 keywords EST.  
 source house mouse.  
 organism Mus musculus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 reference 1 (bases 1 to 310)  
 authors Marra, M., Hillier, L., Allen, M., Bowles, N., Dietrich, N., Dubuque, T.,  
 Gessel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 title The Mashu-HMMI Mouse EST Project  
 journal Unpublished (1996)  
 comment Contact: Marra M/Mouse EST Project  
 Mashu-HMMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LIND; contact the  
 IMAGE Consortium (info@image.lind.gov) for further information.  
 MGI:915085  
 Seq primer: -28m13 rev2 ET from Amersham.  
 features  
 source Location/Qualifiers  
 1..310  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /note="Organ: mammary gland; Vector: pT73D-Pac  
 (Pharmacia) with a modified polylinker; Site.1: Not I -  
 Site.2: Eco RI. 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer [5,  
 TGTACCAATCTGAAGTGGAGCGGCCGCAATGTTTCTTTTCTTTTCTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 RNA provided by Dr. Minoru Ko, Wayne State Univ. Library  
 constructed and normalized by Bento Soares and M. Fatima  
 Bonaldo."  
 /db\_xref="taxon:10090"  
 /clone="1431017"  
 /clone\_lib="Soares mouse mammary gland NBMG"  
 /sex="male"  
 /tissue\_type="mammary gland"  
 /dev\_stage="4 weeks"  
 /lab\_host="DH10B"  
 BASE COUNT 80 a 77 c 85 g 67 t 1 others  
 ORIGIN  
 Query Match 20.8%; Score 149.8; DB 30; Length 310;  
 Best Local Similarity 72.5%; Pred. No. 1.2e-36;  
 Matches 208; Conservative 0; Mismatches 77; Indels 2; Gaps 1;

QY 598 CAGCTCAGCAGCCTGACATCTGAGACACTGCCGTATTACTGTGC 644  
 |||||  
 DB 242 CAGCTCAGCAGCCTGACATCTGAGACACTGCCGTATTACTGTGC 288  
 RESULT 4  
 A1201426/c 461 bp mRNA EST 30-OCT-1998  
 LOCUS A1201426  
 DEFINITION g573c06.x1 NCI CGAP.Pr28 Homo sapiens cDNA clone IMAGE:1943722 3'  
 similar to gb:M8512 IG HEAVY CHAIN PRECURSOR V-1 REGION  
 (HUMAN); contains Alu repetitive element; mRNA sequence.  
 accession A1201426  
 nid g3754032  
 keywords EST.  
 source human.  
 organism Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
 Primates; Catarrhini; Homnidae; Homo.  
 reference 1 (bases 1 to 461)  
 authors NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 title National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 journal Unpublished (1997)  
 comment Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 unknown library type  
 insert length: 608 Std Error: 0.00  
 Seq primer: -40UP from Gldco.  
 features  
 source Location/Qualifiers  
 1..461  
 /organism="Homo sapiens"  
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI-CGAP.Pr28 was prepared, and ss  
 circles were made in vitro. Following HMP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (clones IDs  
 985608-986759, 1101192-1101959, and 1217928-1220615).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1943722"  
 /clone\_lib="NCI-CGAP.Pr28"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 BASE COUNT 93 a 132 c 106 g 130 t  
 ORIGIN

Query Match 20.3%; Score 146.2; DB 31; Length 461;  
 Best Local Similarity 69.3%; Pred. No. 1.8e-35;  
 Matches 199; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
 QY 358 GTGACGCTGACAGCTCTGGGGTTGAGCTTGTCCGAGAGGGGCTTATGTCAGTTGTC 417  
 |||||  
 DB 435 GTGACGCTGCTGACAGCTCTGGGGTTGAGCTTGTCCGAGAGGGGCTTATGTCAGTTGTC 376  
 QY 418 TCCAAAGCTTCGACTTCAACATTAAGACTATTATCCACTGGGTGAAGAGGGCT 477  
 |||||  
 DB 375 TCCAAAGCTTCGACTTCAACATTAAGACTATTATCCACTGGGTGAAGAGGGCT 316  
 QY 478 GAACAGGGCTGGAATGATGATGATGATTCCTGAGAAATGTAATCTGATATGAC 537  
 |||||  
 DB 315 GGACAGAGGCTTGGGTGATGATGATGATGATTCCTGAGAAATGTAATCTGATATGAC 256  
 QY 538 CCGAAATTCAGGGCAGAGCCAGTATTAACAGACACATCTCCACAGCGGCTTACTT 597  
 |||||  
 DB 255 CAGAAGTTTCAGGGCAGAGCCAGTATTAACAGACACATCTCCACAGCGGCTTACTT 196

QY 598 CAGCTAGACGCTGACATCTGAGACACTGCCGCTATTACTGTCC 644  
LOCUS 195 GAGCTGAGCAGCTGAGATCTGAGACAGCGCCAGCTATTACTGTCC 149  
RESULT 5  
AA569186 387 bp mRNA EST 09-SEP-1997  
DEFINITION nm30d10.s1 NCI-CGAP\_L1p2 Homo sapiens cDNA clone IMAGE:1061683  
similar to gb:102325 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN);,  
RNA sequence.  
ACCESSION AA569186  
KEYWORDS g2342240  
NID EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 387)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bbrp/image/image.html

FEATURES  
source  
Insert Length: 1722 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 284.  
Location/Qualifiers  
1..387  
/organism="Homo sapiens"  
/note="Vector: pAMP10; mRNA made from liposarcoma, cDNA  
made by oligo-dT priming. Non-directionally cloned.  
Size-selected on agarose gel, average insert size 600 bp.  
Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."  
/db\_xref="taxon:9606"  
/clone\_image="IMAGE:1061683"  
/clone\_lib="NCI-CGAP\_L1p2"  
/tissue\_type="liposarcoma"  
/lab\_host="DH10B"  
<1..>387

BASE COUNT 89 a 95 c 118 g 85 t  
ORIGIN  
mRNA  
Query Match 19.0%; Score 136.6; DB 22; Length 387;  
Best Local Similarity 67.2%; Pred. No. 1.7e-32;  
Matches 193; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 358 GTGCAGCTCAGCAGCTGGGGTTGAGCTTGCCGAGAGGGCCCTAGTCAAGTTGCC 417  
DB 98 GTGCAGCTGCTGAGCTGGGGCTGAGTGAAGAGCCCTGAGGAGCTTCC 157  
QY 418 TCGAAGCTTCTGACTTCAACATTAAGACTATTATATCCACTGGGTGAAGCAGGCCCT 477  
DB 158 TGTAAAGCATCTGAGATCCACCCCTACCAACTGCCATATGCTGGCAGAGGCCCT 217  
QY 478 GAACAGAGGCTGAGATGATGATTCATCCTGAGAGTGTATATCTGATATGAC 537  
DB 218 GGGCAAGGCGCTGAGTGGTGGGAATGATCAATTTCTAGTGTGTTATATAGTAAGCA 277

QY 538 CCGAATTCAGAGGCGAGGCCAGATTAACAGACACATCTCCACGCGGCTACCTT 597  
DB 278 CAGAGTTCCAGGCGAGAGTACCATGACACAGAGCTCCAGAGCAGTATATG 337  
QY 598 CAGCTCAGAGCGCTGACATCTGAGCAGCTGCCGCTTACTGTGCC 644  
DB 338 GAGCTTAGCAGCTGAGATCTGAGGACAGCGCCGCTATTACTGTGC 384  
RESULT 6  
A1007196 294 bp mRNA EST 12-JUN-1998  
LOCUS A1007196  
DEFINITION ua73904.r1 Soares 2NbMT Mus musculus cDNA clone 1363158.5' similar  
to gb:M18512 IG HEAVY CHAIN PRECURSOR V-I REGION (HUMAN); gb:M12376  
Mouse immunoglobulin H-chain V-region pseudogene mRNA, complete  
(MOUSE);, mRNA sequence.  
ACCESSION A1007196  
KEYWORDS g3216753  
NID EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 294)  
AUTHORS Maira, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.  
TITLE The WashU-HMNI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Maria M/Mouse EST Project  
WashU-HMNI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:896378  
Trace considered overall poor quality  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 1.  
Location/Qualifiers  
1..294  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through two  
rounds of normalization, and was constructed by Bento  
Soares and M.Fatima Bonaldo."

FEATURES  
source  
1..294  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through two  
rounds of normalization, and was constructed by Bento  
Soares and M.Fatima Bonaldo."

BASE COUNT 85 a 68 c 77 g 64 t  
ORIGIN  
Query Match 18.6%; Score 133.8; DB 28; Length 294;  
Best Local Similarity 70.9%; Pred. No. 1.2e-31;  
Matches 205; Conservative 0; Mismatches 82; Indels 2; Gaps 2;

OY		358	GTCGACCTGCAGCAGCTCGGGGTGAG-C-TTGCCAGAGAGGGGCGTTATGTCAGTTC	416
Db		2	GTCCACTGACGACCAGCTGGGGCTGAACATGGTAGAAGTGGGTCTTCAGTAAGCTGTA	61
OY		417	CTCGAAAGCTTCGTCACTTCACATTAAAGAATCATTTATATCCATGGGTGAAGCAGAGCC	476
Db		62	CTCGAAGATTCAGGCTACACCTTCACCGACTACATGATGATCATGGGTGAAGCAGAGCA	121
OY		477	TGACAAGGCCCTGGAAATGATGATGATGATGATCATCTGTGAATGTAATATCTATATGA	536
Db		122	TATAACAAGGCATTTGATGATGATGTAACATTTGACCTTCAGATGAGTAACATCACATCAA	181
OY		537	CCCCAAATTCAGGGCCAGGCCCAGTAAACAGAGACACATCTCCAAAGGGGCGCTAAC-C	595
Db		182	TCAAAAGTTCANAGACNAGGCCCATTTGACTGTAGACANAATTCACACGACACAGCTTACG	241
OY		596	TTCACTGACGACGACCTGACATCTGAGGACATGCCGCTCATTAATCTAGTCGC	644
Db		242	TGCAGCTCGACGAGACTGACATCTGAGGACATCGCGGTCTATATCTGTGC	290
RESULT		7		
LOCUS		AA456778	400 bp	mRNA EST 06-JUN-1997
DEFINITION		Zw7fll.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 770541		
FEATURES		5' similar to gb:X06764 IG KAPPA CHAIN PRECURSOR V-II REGION		
KEYWORDS		(HUMAN); mRNA sequence.		
SOURCE		AA456778		
ORGANISM		human.		
REFERENCE		Homo sapiens		
AUTHORS		Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;		
JOURNAL		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae;		
COMMENT		Unpublished (1997)		
CONTACT:		Contact: Wilson RK		
		WashU-Merck EST Project		
		Washington University School of Medicine		
		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
		Tel.: 314 286 1800		
		Fax: 314 286 1810		
		Email: est@watson.wustl.edu		
		This clone is available royalty-free through LML ; contact the		
		IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information.		
		Trace considered overall poor quality		
		Seq primer: -28mJ rev2 ET from Amerisham		
		High quality sequence stop: 1.		
		Location/Qualifiers		
		1..400		
		/organism="Homo sapiens"		
		/note="Organ: ovary; Vector: pRT73D (Pharmacia) with a		
		modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st		
		strand cDNA was primed with a Not I - oligo(dt) primer [5'		
		TGTTTCACATCTGAGTGAGTGAGCGCGCGGTTTTTTTTTTTTT 3'] ,		
		double-stranded cDNA was size selected, ligated to Eco RI		
		adapters (Pharmacia) , digested with Not I and cloned into		
		the Not I and Eco RI sites of a modified pRT73 vector		
		(Pharmacia). Library constructed by Bento Soares and		
		M.Fatima Bonaldo."		
		/db_xref="taxon:9606"		
		/clone="770541"		
		/clone.lib="Soares ovary tumor NbHOT"		
		/sex="Female"		
		/tissue_type="ovarian tumor"		

mRNA	/lab_host="DH10B (ampicillin resistant)"
<1..>	<.db_xref="GDB:5980414"
BASE COUNT	90 a 120 c 106 g 84 t
ORIGIN	
Query Match	17.1%; Score 122.8; DB 20; Length 400;
Best Local Similarity	65.2%; Pred. No. 3.7e-28;
Matches 214; Conservative	0; Mismatches 107; Indels 7; Gaps 2
OY	1 ATGCAGCTGACCAGTGTCCACGATCATGTCGTGATCCTCCAGGGGAAGAAGTCAACATG 60 
Dd	61 ATTGGTTGAGGACAGCTCTCCAGCACCTTGCTTTGTCTCCAGGGGAAGAAGACCTTC 120 
OY	61 ACCTGACTGCCAGCTCAAGTGAAGTACA-----TGCAGTGGTATCAGCAGAATCA 114 
Dd	121 TCCTGCAGGGCGACGACAGTAGTGTAGCAGACAGCTGCTTAGCCTGGTACAGAGAAACCT 180 
OY	115 AGCAGCTCCCCAAGTCTGGGTTATTATGACACATCCAAACTGGCTTCTGGAGTCCAGGT 174 
Dd	181 GGCCAGGCTCCAGAGCTCCTCATCTATGTTGGTGCATCAGAGGCAACTGGCATCCAGAC 240 
OY	175 GCCTTCAGTGGCAGAGGAGGCTCGAAACGCTTCTCCTCAGCATAGC-AGCATGGAGAGC 233 
Dd	241 AGCTTCAGTGGCAGAGTGGCTTGAGACAGACCTTCACTCTCAACATGGCAGAGCTGAGCC 300 
OY	234 TGAAGATGCTGCACATTTATTTATTTATTCATCAGGGAGTGGGTACCCATTCAGCTTCGCTC 293 
Dd	301 TGAGCATTTTGGAGTGTATTAAGTGTGACAGAGATGATGATGATGATGATGATGATGATGAT 360 
OY	294 GGGGCAAAAGTTGGCAATAAAGTTCT 321 
Dd	361 AGGACCCAGGTGGAAATCAAACGACT 388 
RESULT	8
LOCUS	T27593 393 bp mRNA EST 06-SEP-1995
DEFINITION	EST100653 Homo sapiens CDNA 5' end similar to immunoglobulin kappa light chain, V region (GB:I01279) (HR:3043).
ACCESSION	T27593
NID	9609691
KEYWORDS	EST.
SOURCE	human primer-M13 Reverse library-Human Pancreas.
ORGANISM	Homo sapiens
REFERENCE	Eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Deuterostomia; Chonata; Reptopoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 393) Adams,M.D., Kerlavage,A.R., Fleischman,R.D., Fulder,R.A., Bult,C.J., Lee,N., Kirsch,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A., Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., FitzGerard,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodex,A., Grehm,C.L., Hanna,M.C., Hedblom,E., Hinkle Jr,P.S., Kelley,J.M., Klimke,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palancues,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J., Dimke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.-W., Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L., Kunsch,C., Ji,H., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wel,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C. Initial Assessment of Human Gene Diversity and Expression Patterns Based Upon 52 Million Basepairs of cDNA Sequence Unpublished (1995)
JOURNAL COMMENT	Other_ESTs: THC24356 Contact: Venter, JC

The Institute for Genomic Research  
9332 Clopper Rd, Gaithersburg, MD 20878  
Tel: 301.861.9056  
Fax: 301.861.9423  
Email: [tdbinfo@tdb.tigr.org](mailto:tdbinfo@tdb.tigr.org)  
For clone availability, additional sequence and expression  
information related to this EST, please contact the TIGR Database  
[tdbinfo@tdb.tigr.org](mailto:tdbinfo@tdb.tigr.org).

FEATURES	location/Qualifiers
SOURCE	1..393
MRNA	/organism="Homo sapiens"
BASE COUNT	<1..>393
ORIGIN	88 a 113 c 96 g 94 t 2 others

Query Match	17.0%;	Score 122.2;	DB 11;	Length 393;
Best Local Similarity	64.9%;	Pred. No. 5.7e-28;		
Matches 211; Conservative	0;	Mismatches 110;	Indels 4;	Gaps 2

[illegible]

RESULT	9		
AA295311			
LOCUS	328 bp	mRNA	EST
DEFINITION	EST100471 Pancreas tumor I Homo sapiens cDNA 5' end similar		18-APR-1998
	immunoglobulin kappa light chain, VJ regions, mRNA sequence		
ACCESSION	AA295311		
NID	g1947646		
KEYWORDS	EST.		
SOURCE	human.		

REFERENCE  
AUTHORS

1 (bases 1 to 328)

Adams,M.D., Kerlavange,A.R., Fleischmann,R.D., Fulcher,R.A.,  
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,  
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wal,C.,  
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,  
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.,  
Glodet,C., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,  
Kelley,J.M., Kelley,J.C., Liu,L.-T., Marmaros,S.M., Merick,J.M.,  
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,  
Small,K.V., Spriggs,T.A., Utteback,T.R., Weidman,J.F., Yi,Y.,  
Bednarzik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dilke,D., Feng,D.-F., Fertle,A., Fischer,C., Hastings,G.A.,  
He,M.W., Hu,U.S., Greene,J.M., Gruber,T., Hudson,P., Kim,A.K.,  
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,

TITLE	Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL	Nature 377 (6547 Suppl.), 3-174 (1995)
MEDLINE	96026280
COMMENT	Other-ESTs: TRC169106

FEATURES	source	Location/Qualifiers
		1..328
		/organism="Homo sapiens"
		/note="Organ: pancreas; Vector: pbluescript SK-; Site:1 EcoRI; Site:2 XhoI"
		/db_xref="ATCC (inhost):190766"
		/db_xref="taxon:9606"
		/clone_lib="Pancreas tumor I"
		/dev_stage="adult"
mRNA		<1..>328
BASE COUNT	73 a 93 c 78 g 77 t 7 others	
ORIGIN		

Query Match	16.5%	Score 119;	DB 17;	Length 328;
Best Local Similarity	65.0%;	Pred. NO. 5.4e-27;		
Matches 186; Conservative	0;	Mismatches 97;	Indels 3;	Gaps 1.

QY	1	ATGCAGTGCACCCAGTGTCCAGCAGATATGTCTGCATCTCCAGGGGAAAGATCCACATG	60
Db	43	ATACTGATGACGACAGNTTNCAGCCACCCTGTCTGTGTCTCAGGGGAAAGACCCACTTN	1020
QY	61	ACCTGCAGTGCACGCTCAAGTGTG -- AGTAACTGCATCGTGGTATCAGCAGAAATCAAGC	117
Db	103	TCTCGCAGGGGCACATCGATCAGAGTGTAGCAGCAACTTACCTGGTATCCAGCAGAAACTTGGC	1670
QY	118	ACCTGCCCAACACCTGTGGGTTATGACACATCCAAATCGCTTCTGGAGTCCAGCTGC	177
Db	163	CAGGCTCCCAAGGCTCTTCATCTATTAGGTGCATTCATCAGGGGNCATCTGGCATCTCCAGCCAGG	222
QY	178	TTTCAGTGGCAGTGGGCTGTGCAAACTCTTACTCTCTCAAGATTCAGCAGACATGGAGGCTGAA	237
Db	223	TTTCAGTGGCAGTGGGCTGTGGACAGAGTTCACCTCTCANCAATCAGACATCTGCAGGTGAA	287
QY	238	GATGCTGCACACTTATTTATGTATCAGGGGAGTGGGTATCCCATTTCA	283
Db	283	GATTTTNCAGTTTATTTACTGTACACCAATATAATAATCAGGGCATTTCA	328

RESULT	10
LOCUS	AA710291
DEFINITION	AA710291 345 bp mRNA EST 24-DEC-1997
	v53a04.1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
	clone 1166766 5' similar to gb:X02484-rna1 IG KAPPA CHAIN PRECURSOR
	V-V REGION (HUMAN): gb:J00560 mouse Ig kappa mRNA from mopc21 &
	other myeloma mRNA 3' (MOUSE), mRNA sequence.
ACCESSION	AA710291
NID	g2720209
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus.
	Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
	Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 345)



QY	154	162	QY	214	Db	222	QY	274	Db	282	RESULT 12	AA295093	LOCUS	DEFINITION	ACCESSION	AA295093	IID	91947582	KEYWORDS	EST.	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT	FEATURES	source																						
	154	162	QY	214	Db	222	QY	274	Db	282	AA295093	370 bp	MRNA	EST	18-Apr-1997	AA295093	EST100400	Pancreas tumor I Homo sapiens cDNA 5' end similar to similar to immunoglobulin kappa, variable region (GB:Y00640), mRNA sequence.	AA295093	91947582	EST.	human.	Homo sapiens	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	1 (bases 1 to 370)	Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Well,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fline,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fitchman,J.L., Georgagen,N.S., Glodetz,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.W., Kelsey,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R., Small,K.V., Spilggs,T.A., Uteback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Gao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dinke,D., Feng,D.-F., Fertile,A., Fischer,C., Hastings,G.A., He,W.M., Hu,J.S., Greene,J.M., Gruber,J., Hudson,J., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weisner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.	Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence Nature 377 (6547 Suppl.), 3-174 (1995)	96026280	Other ESTs: THC168243	Contact: Kerlavage, AR	Bioinformatics	The Institute for Genomic Research	9712 Medical Center Drive, Rockville, MD 20850 USA	Tel: 3018699056	Fax: 3018699423	Email: arkerlav@tigr.org	For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index ( <a href="http://www.tigr.org/tdb/hgi/hgi1.html">http://www.tigr.org/tdb/hgi/hgi1.html</a> )	Seq primer: M13 Reverse.	Location/Qualifiers	1. 370	/organism="Homo sapiens"	/note="Organ: pancreas; Vector: pBluescript SK-; Site:1: ECORI, Site:2: XhoI"	/db_xref="ATCC (inhost):190705"	/db_xref="taxon:9606"	/clone_lib="Pancreas tumor I"	/dev_stage="adult"	<1. ->370	85 a	107 c	94 g	77 t	7 others

ORIGIN	Query Match	16.3%	Score 117.6	DB 17	Length 370
	Best Local Similarity	66.9%	Pred. No. 1.5e-26		
	Matches 178: Conservative	0	Mismatches 85	Indels 3	Gaps 1
OY	1 ATGCAGCTGACCCAGCTCCAGCAATCATGTGTCATCCAGGGGAAAGGTCCACCATG 60				
Db	50 ATAGTATGATGAGCGCANNNTCCAGCACCTGCTGCTGTGTCCAGGGGAAAGGCCACCTTC 109				
OY	61 ACCTGCACTGCCAGCTCAAGTGTAAAGT--AACATGCAGCTGTATCAGCAGAACTCAAGC 117				
Db	110 TCCTCCAGAGGCGCAGTACAGAGTGTATGAGCAACTAGCTGCTGTACACAGAAACTCTGGC 169				
OY	118 ACCTCCCCCAACTCTGGCGTTATATGACACATTCACAACTGGCTTCTGGAGTCCAGGTGCG 177				
Db	170 CAGGACCCAGGAGCTCTCATATCATATCATCCACCAAGGCGCAGCTGTATCCACCCAGG 229				
OY	178 TTCAGTGGCAGTGGGTGCGAAACCTTACGTCCTCAGATCAGACAGATGAGAGCTGAA 237				
Db	230 TTCACTGGCAGTGGGTGCGAGAGTTCACCTACCTACATCAGCAGCTACAGTCTGAA 289				
OY	238 GATGCTGCCACTTATTTATTTATTTATCA 263				
Db	290 GATTCTGACGTTTATTTACTGTCAGCA 315				
RESULT 13					
AA300732	333 bp	mRNA	EST	18-Apr-1997	
LOCUS	EST13847	Testis tumor Homo sapiens CDNA 5' end similar to similar to immunoglobulin kappa light chain (GB:S4906), mRNA sequence.			
DEFINITION	AA300732				
ACCESSION	g1953300				
NID	EST.				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1 (bases 1 to 333)				
AUTHORS	Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstein,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-vel,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fife,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Georgagen,N.S., Glodet,A., Gnehm,C.L., Hanna,M.C., Hedlorn,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudak,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y., Beharavik,D.P., Cao,L., Cepede,M.A., Coleman,T.A., Collins,E.J., Dinke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.M., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kodak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.				
TITLE	Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence				
JOURNAL	Nature 377 (6547 Suppl), 3-14 (1995)				
MEDLINE	96026280				
COMMENT	Other ESTs: THC87411 Contact: Kerlavage, AR Bioinformatics The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850 USA Tel: 3016699056 Fax: 3016699423 Email: arkerlav@etlgr.org For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene				



Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse

DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

Insert length: 1070 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amerisham  
High quality sequence stop: 413.  
Location/Qualifiers

## FEATURES

source  
1..418  
/organism="Homo sapiens"  
/note="Vector: pAMP10; mRNA made from liposarcoma, CDNA  
made by oligo-dt priming. Non-directionally cloned.  
Size-selected on agarose gel, average insert size 600 bp.  
Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."  
/db\_xref="taxon:9606"  
/clone="IMAGE:940044"  
/clone\_1lb="NCI\_CGAP\_L1P2"  
/tissue\_type="liposarcoma"  
/lab\_host="DH10B"  
<1..>418

MRNA  
BASE COUNT 99 a 127 c 102 g 90 t  
ORIGIN

## Query Match

16.0%; Score 115.4; DB 21; Length 418;

Best Local Similarity 63.9%; Pred. No. 7.8e-26;  
Matches 209; Conservative 0; Mismatches 111; Indels 7; Gaps 2;

```
QY 1 ATGCACTGACCCAGTCTCCAGCAATCATCTGCAATCTCCAGGGGAAAAGTCCACCATG 60
   || || || || || || || || || || || || || || || || || || || || ||
DB 86 ATTGTGTGACGAGCTCTCCAGCACCCTGTTGTCTCCAGGGGAAAAGGCCACCTTC 145
   || || || || || || || || || || || || || || || || || || || || ||
QY 61 ACTGCAGTGCAGCTCAAGTGAAGTAAATGCA-----CTGGTATCAGCAGAGTCA 114
   || || || || || || || || || || || || || || || || || || || || ||
DB 146 TCTTGCAAGGGCCAGTCAAGTATCATCAGCACCCTTAGCCTGGTATCAGCAAAACT 205
   || || || || || || || || || || || || || || || || || || || || ||
QY 115 AGCACCTCCCCCAACTCTGGTTTATGACACATCCAACTGSGCTTGGAGTCCAGGT 174
   || || || || || || || || || || || || || || || || || || || || ||
DB 206 GGGCAGAGCTCCAGGCTCTCATCTACGGTGCATGCAGAGGCCACTGGA--TCCAGAC 264
   || || || || || || || || || || || || || || || || || || || || ||
QY 175 CGCTTCAGTGGCAGTGGGTCTGGAACCTTACTCTCTCAGCATCAGCAGCATGAGGCT 234
   || || || || || || || || || || || || || || || || || || || || ||
DB 265 AGGTCAGTGGCAATTCGTCTGGACAGACTTCAGTCTCACCATCGCAGACTGGAGCCT 324
   || || || || || || || || || || || || || || || || || || || || ||
QY 235 GAAGATGCTGCCACTATATTGTTATTCAGGGGAGTGGGTAACCATTCACGTTCCGCTCG 294
   || || || || || || || || || || || || || || || || || || || || ||
DB 325 GAAGATTCTGCACTTACTATTGTCAGCACTATGTAACCTCATTTATGACATTCGGCCAG 384
   || || || || || || || || || || || || || || || || || || || || ||
QY 295 GGGACAAAGTTGGAAATAAAGGTTCT 321
   || || || || || || || || || || || || || || || || || || || || ||
DB 385 GGGACCAAGTGGAAATCAACGAACT 411
   || || || || || || || || || || || || || || || || || || || || ||
```

Search completed: May 13, 1999, 09:40:44  
Job time: 362 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: May 13, 1999, 09:55:24 ; Search time 59.56 Seconds  
(without alignments)  
2274.252 Million cell updates/sec

Title: US-08-704-178-2

Perfect score: 720  
Sequence: 1 ATGCAGCTGACCCAGTCTCC.....CGGTCCAGCTCTCTCATTA 720

Scoring table: IDENTITY\_NUC

Searched: 240622 seqs, 94065609 residues

Database : N.Geneseq\_34.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	720	100.0	720	1	Sequence encoding
2	720	100.0	720	1	Single-chain anti-
3	433.4	60.2	711	1	Anti-erbB2 scfv CD
4	433	60.1	711	1	Single-chain anti-
5	432	60.0	732	1	Fv(gp-2) immunosp
6	429.8	59.7	711	1	Sequence encoding
7	404.4	56.2	738	1	OV81 light and hea
8	383.8	53.3	720	1	CDNA encoding an a
9	326.4	45.3	735	1	CDNA encoding an a
10	322	44.7	720	1	Coding sequence of
11	322	44.7	720	1	Single chain bindi
12	322	44.7	720	1	18-2-3/TRY202' sin
13	321.6	44.7	724	1	18-2-3/TRY202' S
14	321.2	44.6	741	1	Coding sequence of
15	321.2	44.6	741	1	Single chain bindi
16	321.2	44.6	741	1	TRY59 single chain
17	319.6	44.4	741	1	TRY59. Single poly
18	317.2	44.1	753	1	TRY104b. Single po
19	317.2	44.1	753	1	Coding sequence of
20	317.2	44.1	753	1	Single chain bindi
21	317.2	44.1	753	1	TRY104b single cha
22	312.6	43.4	906	1	Anti-human SC sing
23	310.2	43.1	729	1	Coding sequence of
24	310.2	43.1	729	1	Single chain bindi
25	310.2	43.1	729	1	18-2-3/TRY59. Sin
26	307	42.6	729	1	Nucleic acid seque
27	300.8	41.8	6727	1	Nucleic acid seque
28	299.6	41.6	1320	1	Nucleic acid seque
29	299.6	41.6	6799	1	Nucleic acid seque
30	299.6	41.6	1299	1	Nucleic acid seque
31	299	41.5	729	1	Fv(gp-4) immunosp
32	298.6	41.5	1679	1	Single chain anti-
33	298.6	41.5	1679	1	H2-anti-CEA anti-
34	293.6	40.8	1460	1	Sequence encoding
35	293.6	40.8	803	1	PLAP CC49/212 SCA
36	292.6	40.6	797	1	CC49/212 SCA PLAP
37	290.6	40.4	731	1	Sequence encoding
38	289	40.1	819	1	Sequence encoding
39	288.8	40.1	861	1	Sequence encoding
40	287.8	40.0	711	1	TRF61. Single poly
41	287.8	40.0	711	1	Coding sequence of
42	287.8	40.0	711	1	Single chain bindi
43	287.8	40.0	711	1	TRY61 single chain

ALIGNMENTS

44	285.4	39.6	1065	1	T94963
45	285.4	39.6	1137	1	T94964
ALIGNMENTS					
RESULT	1				
ID	05181				
AC	05181				
DE	21-JUN-1994 (first entry)				
DE	Sequence encoding the single chain anti-erbB2 antibody, Ab no.21.				
KW	Single chain anti-erbB2 antibody; cancer therapy; prevention;				
OS	Synthetic.				
FS	Key	Location/Qualifiers			
FT	cds	1..171			
FT	cds	/*tag- a			
PN	WO9400136-A.				
PD	06-JAN-1994.				
PF	21-OCT-1992; U08545.				
PR	30-JUN-1992; US-906555.				
PA	(MOLE-) MOLECULAR ONCOLOGY INC.				
PI	Kasprzyk Fg, King CR;				
DR	P-PSDB; R45443.				
PT	Treatment of malignancies over-expressing ERB-12 - using at least				
PT	2 monoclonal antibodies which recognise different epitopes on				
PT	gp185				
PS	Example: Fig 8; 37pp: English.				
CC	The source of human erbB-2 protein for the prodn. of antibodies no.				
CC	23 and 21 (05180/R45442; 05181/R45443) is a NIH/3T3 cell				
CC	engineered to express the human erbB-2 protein on its surface (N/				
CC	erbB-2). Abs no. 21 and 23 are directed against the extracellular				
CC	domain of gp185 erbB-2. Nude mice manipulated to produce rapidly				
CC	growing tumours were used in a trial of the efficacy of the Abs. In				
CC	animals given a combination of the 2 Abs, tumours completely				
CC	regressed after 11 days.				
SO	Sequence 720 BP; 184 A; 178 C; 179 G; 179 T;				
Query Match					
Best Local Similarity 100.0%; Score 720; DB 1; Length 720;					
Matches 720; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
OY	1	ATGCAGCTGACCCAGTCTCTCATCATGTCATCTCCAGGGGAAAGTCCACG	60		
DB	1	ATGCAGCTGACCCAGTCTCTCATCATGTCATCTCCAGGGGAAAGTCCACG	60		
OY	61	ACCTGAGTGCAGGCAAGTGAATGATGATGATGATGATGATGATGATGATG	120		
DB	61	ACCTGAGTGCAGGCAAGTGAATGATGATGATGATGATGATGATGATGATG	120		
OY	121	TCGCCAAACTCTGGTTTATGACACATCCAACTGGCTTCGAGTCCAGGCTTC	180		
DB	121	TCGCCAAACTCTGGTTTATGACACATCCAACTGGCTTCGAGTCCAGGCTTC	180		
OY	181	AGTGGAGTGGGTCTGGAACCTTACTCTCAGCATAGAGCATGAGGCTGAAGT	240		
DB	181	AGTGGAGTGGGTCTGGAACCTTACTCTCAGCATAGAGCATGAGGCTGAAGT	240		
OY	241	GGTGCACCTTATTTATGATATGATGATGATGATGATGATGATGATGATG	300		
DB	241	GGTGCACCTTATTTATGATATGATGATGATGATGATGATGATGATGATG	300		
OY	301	AACTGGAATTAAGGTTCTACCTCCGATCTGTTAACTCTGTAAGTAAAGTGTG	360		
DB	301	AACTGGAATTAAGGTTCTACCTCCGATCTGTTAACTCTGTAAGTAAAGTGTG	360		
OY	361	CAGCTCAGCAGTCTGGGGTTGAGCTTGCAGAGAGGGCCCTTACAGTTGCTCC	420		
DB	361	CAGCTCAGCAGTCTGGGGTTGAGCTTGCAGAGAGGGCCCTTACAGTTGCTCC	420		

QY	421	AAAGCTTGACCTTCAACATTAAGAAGCATATATATCCACAGGGGTGAAGCAGAGGCCGAA	480
Db	421	AAAGCTTGACCTTCAACATTAAGAAGCATATATATCCACAGGGGTGAAGCAGAGGCCGAA	480
QY	481	CAGGGCCCTGAGATGAGTATGATGATTCATCCCTGAGATGGTAACTGTATATGACCCG	540
Db	481	CAGGGCCCTGAGATGAGTATGATGATTCATCCCTGAGATGGTAACTGTATATGACCCG	540
QY	541	AAATTCGAGGCAAGGCCAGATTAACAAGCAGACATCTCCAAAGGGGGCTACCTTCAG	600
Db	541	AAATTCGAGGCAAGGCCAGATTAACAAGCAGACATCTCCAAAGGGGGCTACCTTCAG	600
QY	601	CTCAGACACCTGACATCTGAGAGCACTGCCGCTCTTACTCTGTCTTTATACTACTAT	660
Db	601	CTCAGACACCTGACATCTGAGAGCACTGCCGCTCTTACTCTGTCTTTATACTACTAT	660
QY	661	AGTCTTACTATAGCTATGATCTACTACTGGGGGTAAAGAAACCTTCGACACCGTCTCTATA	720
Db	661	AGTCTTACTATAGCTATGATCTACTACTGGGGGTAAAGAAACCTTCGACACCGTCTCTATA	720

RESULT	2
T65007	
ID	T65007 standard; cDNA; 720 BP.

DE 05-JUN-1997 (first entry)  
 DT Single-chain anti-erbB2 antibody e21(Fv) cDNA.  
 KM Single chain antibody; variable region; light chain; heavy chain.  
 KW breast cancer; ovarian cancer; non-small cell lung carcinoma;  
 KM immunodiagnostics; treatment; cytotoxic agent; erbB-2; ds.  
 OS Mus musculus.  
 DS Synthetic.

	Key	Location/Qualifiers
FH		1. .720
FT	mat_peptide	/*tag= a
FT		/product= e21(Fv)
FT		

PN US587458-A.  
PD 24-DEC-1996.  
PE 07-OCT-1991. 772270.  
PR 07-OCT-1991. 772270.  
PR 30-JUN-1992. US-906555.  
PR 14-MAY-1993. US-061092.  
PR (AIRON-) AROHEX PHARM INC.  
PA BILD RE, KSPPTZ Pg, KING CR,  
PI WPI: 97-064831/06.

PT Single chain antibodies specific for erbB-2 protein, gp185  
PT labels or cytotoxin, useful for detection and treatment of tumour  
PT cells expressing this protein  
PS Example 9; Columns 27-30; 28pp; English.

CC The present cDNA sequence codes for a claimed single-chain antibody  
CC designated e21(Fv), which binds to erbB-2. Monoclonal antibody e21  
CC was generated by immunising mice with N/erbB-2 cells overexpressing  
CC the gp185 protein, removing spleen cells and producing hybridomas  
CC by standard techniques. Messenger RNA coding for the anti-erbB-2  
CC monoclonal antibody was isolated and converted to cDNA. Regions  
CC coding for the heavy- and light- chain variable regions were then  
CC amplified by PCR and joined via a sequence encoding a peptide  
CC linker. The resulting single-chain antibody is useful for *in vitro*  
CC diagnosis of tumour cells which overexpress the erbB-2 gp185  
CC marker, e.g. breast, ovarian and non-small cell lung carcinomas,  
CC and, when coupled to a cytotoxic agent, to treat such tumours.  
SQ Sequence 720 BP; 144 A; 178 G; 179 T;

Query Match	100.0%	Score 720;	DB 1;	length 720;
Best Local Similarity	100.0%	Pred. No. 2e-182;		
Matches 720; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy.  
1 ATGCAGCTGACCCAGTCTCCAGCAATCATGTGCATCTCCAGGGGGAAAAGTCCACATG 60  
? |||||  
Db 1 ATGCAGCTGACCCAGTCTCCAGCAATCATGTGCATCTCCAGGGGGAAAAGTCCACATG 60

QY	61	ACGCGAGTGGCAGCTCAAGGTAAAGTAAACATGCACGTGATATCAGACGAAATCAACACC	120
Dp	61	ACCTGCAGTGGCAGCTCAAGGTAAAGTAAACATGCACGTGATATCAGACGAAATCAACACC	120
QY	121	TCGCCCAAACTCTGGGTATTATGACACATCCCAACTGGCTCTTGAGTCCAGGCTTC	180
Dp	121	TCGCCCAAACTCTGGGTATTATGACACATCCCAACTGGCTCTTGAGTCCAGGCTTC	180
QY	181	AGTGGCAGTGGGTCTGGAAACTCTTACTCTCTCAGCATCAGCAGCATGAGGCTGAAGAT	240
Dp	181	AGTGGCAGTGGGTCTGGAAACTCTTACTCTCTCAGCATCAGCAGCATGAGGCTGAAGAT	240
QY	241	GCTGCGCACTTTATTTATGTATCAGGGAGTGGGAGACCATTCAGTTCCGCTCGGGGACA	300
Dp	241	GCTGCGCACTTTATTTATGTATCAGGGAGTGGGAGACCATTCAGTTCCGCTCGGGGACA	300
QY	301	AAAGTGGAAATAAAGGTTTACTCCGAGTCTGGTAAATCTTCTGAAAGGTAAAGGTGTG	360
Dp	301	AAAGTGGAAATAAAGGTTTACTCCGAGTCTGGTAAATCTTCTGAAAGGTAAAGGTGTG	360
QY	361	CAGCTGCAGCAGTGTGGGGGTGAGCTGTCTCCGAGAGGGGCTTATGTCAGATTGTCTGC	420
Dp	361	CAGCTGCAGCAGTGTGGGGGTGAGCTGTCTCCGAGAGGGGCTTATGTCAGATTGTCTGC	420
QY	421	AAAGCTTCTGCATTCACACATTTAAAGCATTTATATCCACAGTGGGTGAAGCAGAGGCTGAA	480
Dp	421	AAAGCTTCTGCATTCACACATTTAAAGCATTTATATCCACAGTGGGTGAAGCAGAGGCTGAA	480
QY	481	CAGGCGCTGAGATGGATTGGATGGATTCACTCTGAGATGGTAATACTGTATATGACCGG	540
Dp	481	CAGGCGCTGAGATGGATTGGATGGATTCACTCTGAGATGGTAATACTGTATATGACCGG	540
QY	541	AAATTTCCAGGGCAGGGCAGATTAACAGCAGACACATCTCCAAAGGGGCTACCTTCAG	600
Dp	541	AAATTTCCAGGGCAGGGCAGATTAACAGCAGACACATCTCCAAAGGGGCTACCTTCAG	600
QY	601	CTCAGCAGGCTGACATCTGAGGACACACGCGGTATTAATCACTGTGCTTTATTAATACTAT	660
Dp	601	CTCAGCAGGCTGACATCTGAGGACACACGCGGTATTAATCACTGTGCTTTATTAATACTAT	660
QY	661	AGTGCCTTAATTCGTATGTACTACTTGGGGGTCAAGGAACCTCGGTACACGCTCTCTATATA	720
Dp	661	AGTGCCTTAATTCGTATGTACTACTTGGGGGTCAAGGAACCTCGGTACACGCTCTCTATATA	720

### RESULT 3

T17728  
ID T17728 standard; cDNA; 711 BP

AC T17728;  
DT 21-MAY-1996 (first entry)

DE Anti-erbB2 scFv cDNA.  
KM Oncoprotein; erbB2; cell proliferation; tumour; cancer;  
KW intracellular antibody homologue; single chain antibody; scFv;  
KM gene therapy; ds.

OS	Synthetic.
PN	W09607321-A1.

PD 14-MAR-1996.  
PF 23-AUG-1995;  
00 SEP 1994

PR	06-SEP-1994; US-301339.
PR	06-JUN-1995; US-468252.
PR	(HARD) THE DEC FRONT

FA (OABR-) OAB RES FOUND  
PI Curiel DT, Deshane J,  
DR WBT: 96-1771307/17

DR P-PSDB; R94020.  
PT Inhibition of

cells - by introducing  
which is expressed and

Claim 42; Page 29-30; 48pp; English.

CC (R94019) linked to a second sequence (T17728) encoding a single  
CC chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.

CC The anti-erbB2 sfv portion is obtained by PCR using e23scfv

CC plasmid as template. The signal peptide directs the scfv to the  
CC endoplasmic reticulum. The nucleic acid is incorporated into a  
CC plasmid or viral vector to facilitate expression of the scfv antibody  
CC homologue within e.g. an epithelial carcinoma cell. Intracellular  
CC expression of the homologue inhibits surface expression of erbB2 and  
CC thereby inhibits cell proliferation and cell survival and decreases  
CC tumorigenicity.  
SQ Sequence 711 BP; 175 A; 183 C; 189 G; 164 T;

Query Match 60.2%; Score 433.4; DB 1; Length 711;  
Best Local Similarity 79.6%; Pred. No. 2.2e-106;  
Matches 512; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 2 TGCAGCTGACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAGTCAACATGA 61  
DB 8 TGCAGCTGACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAGTCAACATGA 67  
QY 62 CTTGAGTGGCAGCTCAAGTGAATGAATCAATGCTGATATCAGCAGAAAGTCAACACT 121  
DB 68 CTTGAGTGGCAGCTCAAGTGAATGAATCAATGCTGATATCAGCAGAAAGTCAACACT 127  
QY 122 CCCCCAACTCTGGGTTTATGACATATCAATGCTGAGTCCAGGAGCTTCA 181  
DB 128 CCCCCAACTCTGGGTTTATGACATATCAATGCTGAGTCCAGGAGCTTCA 187  
QY 182 GTGGCAGTGGGTCTGGAACCTCTTACTCTCAGCATCAGCAGATGGAAGTGAAGT 241  
DB 188 GTGGCAGTGGGTCTGGAACCTCTTACTCTCAGCATCAGCAGATGGAAGTGAAGT 247  
QY 242 CTGCCACTTATTTATGTTATCAGGGGAGTGGTACCATTCACTGCTGGGAGCA 301  
DB 248 CTGCCACTTATTTATGTTATCAGGGGAGTGGTACCATTCACTGCTGGGAGCA 307  
QY 302 AGTTGGAATATAAGGTTCTACCTCCGATCGTAATCTTCAAGTAAAGTGTGC 361  
DB 308 AGTTGGAATATAAGGTTCTACCTCCGATCGTAATCTTCAAGTAAAGTGTGC 367  
QY 362 AGCTCAGCAGTCTGGGTTGATGCTTCCAGAGAGGGGCTTATGCAAGTGTCTGCA 421  
DB 368 AGCTCAGCAGTCTGGGTTGATGCTTCCAGAGAGGGGCTTATGCAAGTGTCTGCA 427  
QY 422 AAGCTTCTGACTTCAACTTAAGACTATTAATTCACACTGGGTGAAGAGAGGCTGAA 481  
DB 428 AAGCTTCTGACTTCAACTTAAGACTATTAATTCACACTGGGTGAAGAGAGGCTGAA 487  
QY 482 AGGGCTCGAATGATGATGATGATTCATCTGAGAAATGTAATCTGATATGACCGGA 541  
DB 488 AGGGCTCGAATGATGATGATGATTCATCTGAGAAATGTAATCTGATATGACCGGA 547  
QY 542 AATTCCAGGGCAGGAGGAGTAAACAGACAGACATCTCCAGAGGGGCTTACCTTCA 601  
DB 548 AATTCCAGGGCAGGAGGAGTAAACAGACAGACATCTCCAGAGGGGCTTACCTTCA 607  
QY 602 TCAGCAGCTGACATCTGAGGACACTGCGCTTATTAAGTGTGC 644  
DB 608 TCAGCAGCTGACATCTGAGGACACTGCGCTTATTAAGTGTGC 650

RESULT 4  
T65006 165006 standard; cDNA; 711 BP.  
AC T65006:  
DT 05-JUN-1997 (first entry)  
DE Single-chain anti-erbB2 antibody e23(Fv) cDNA.  
KW Single chain antibody; variable region; light chain; heavy chain;  
KW breast cancer; ovarian cancer; non-small cell lung carcinoma;  
KW immunodiagnosis; treatment; cytotoxic agent; erbB-2; ds.  
OS Mus musculus.  
SS Synthetic.  
FH Key Location/Qualifiers  
FT mat\_peptide 1..711  
FT /\*tag- a

FT US5587458-A. /product- e23(Fv)  
PN 24-DEC-1996.  
PD 07-OCT-1991: 772270.  
PF 07-OCT-1991: US-772270.  
PR 30-JUN-1992; US-906555.  
PR 14-MAY-1993; US-061092.  
PA (ARON-) ARONEX PHARM INC.  
PI Bird RE, Kaspriyik PG, King CR;  
DR P-PSDB: W15185.  
PT Single chain antibodies specific for erbB-2 protein, gp185 - with  
PT labels or cytotoxin, useful for detection and treatment of tumour  
PT cells expressing this protein  
PS Example 8: Columns 25-28; 28pp: English.  
CC The present cDNA sequence codes for a claimed single-chain antibody,  
CC designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23  
CC was generated by immunising mice with N/erbB-2 cells overexpressing  
CC the gp185 protein, removing spleen cells and producing hybridomas  
CC by standard techniques. Messenger RNA coding for the anti-erbB-2  
CC monoclonal antibody was isolated and converted to cDNA. Regions  
CC coding for the heavy- and light-chain variable regions were then  
CC amplified by PCR and joined via a sequence encoding a peptide  
CC linker. The resulting single-chain antibody is useful for in vitro  
CC diagnosis of tumour cells which overexpress the erbB-2 gp185  
CC marker, e.g. breast, ovarian and non-small cell lung carcinomas,  
CC and, when coupled to a cytotoxic agent, to treat such tumours.  
SQ Sequence 711 BP; 175 A; 182 C; 189 G; 164 T;

Query Match 60.1%; Score 433; DB 1; Length 711;  
Best Local Similarity 79.5%; Pred. No. 2.8e-106;  
Matches 511; Conservative 1; Mismatches 131; Indels 0; Gaps 0;

QY 2 TGCAGCTGACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAGTCAACATGA 61  
DB 8 TGCAGCTGACCCAGTCTCCAGCAATCATGTCTGCATCTCCAGGGGAAAAGTCAACATGA 67  
QY 62 CTTGAGTGGCAGCTCAAGTGAATGAATCAATGCTGATATCAGCAGAAAGTCAACACT 121  
DB 68 CTTGAGTGGCAGCTCAAGTGAATGAATCAATGCTGATATCAGCAGAAAGTCAACACT 127  
QY 122 CCCCCAACTCTGGGTTTATGACATATCAATGCTGAGTCCAGGAGCTTCA 181  
DB 128 CCCCCAACTCTGGGTTTATGACATATCAATGCTGAGTCCAGGAGCTTCA 187  
QY 182 GTGGCAGTGGGTCTGGAACCTCTTACTCTCAGCATCAGCAGATGGAAGTGAAGT 241  
DB 188 GTGGCAGTGGGTCTGGAACCTCTTACTCTCAGCATCAGCAGATGGAAGTGAAGT 247  
QY 242 CTGCCACTTATTTATGTTATCAGGGGAGTGGTACCATTCACTGCTGGGAGCA 301  
DB 248 CTGCCACTTATTTATGTTATCAGGGGAGTGGTACCATTCACTGCTGGGAGCA 307  
QY 302 AGTTGGAATATAAGGTTCTACCTCCGATCGTAATCTTCAAGTAAAGTGTGC 361  
DB 308 AGTTGGAATATAAGGTTCTACCTCCGATCGTAATCTTCAAGTAAAGTGTGC 367  
QY 362 AGCTCAGCAGTCTGGGTTGATGCTTCCAGAGAGGGGCTTATGCAAGTGTCTGCA 421  
DB 368 AGCTCAGCAGTCTGGGTTGATGCTTCCAGAGAGGGGCTTATGCAAGTGTCTGCA 427  
QY 422 AAGCTTCTGACTTCAACTTAAGACTATTAATTCACACTGGGTGAAGAGAGGCTGAA 481  
DB 428 AAGCTTCTGACTTCAACTTAAGACTATTAATTCACACTGGGTGAAGAGAGGCTGAA 487  
QY 482 AGGGCTCGAATGATGATGATGATTCATCTGAGAAATGTAATCTGATATGACCGGA 541  
DB 488 AGGGCTCGAATGATGATGATGATTCATCTGAGAAATGTAATCTGATATGACCGGA 547  
QY 542 AATTCCAGGGCAGGAGGAGTAAACAGACAGACATCTCCAGAGGGGCTTACCTTCA 601  
DB 548 AATTCCAGGGCAGGAGGAGTAAACAGACAGACATCTCCAGAGGGGCTTACCTTCA 607

QY 602 TCAGCAGCTGACATCTGAGACACTCCGCTTACTCTGTC 644  
 DB 608 TCCTCAGTCTGACATCTGAGACTCTGCAGTCTTACTGTC 650

## RESULT 5

QY 073678  
 ID Q73678 standard; DNA; 732 BP.  
 AC Q73678;  
 DT 09-MAY-1995 (first entry)  
 DE Fv(GP-2) immunosuppressive.  
 KM Mab; monoclonal antibody; hybridoma; interleukin-2; IL-2;  
 KW Fv; antibody variable region; GP-2; Fv(GP-2);  
 OS immunosuppressive; ss.  
 PN Mus sp.  
 PD EP-621338-A.  
 PE 26-OCT-1994.  
 PF 21-APR-1994; 106257.  
 PR 21-APR-1993; JP-094491.  
 PR 07-MAR-1994; JP-036065.  
 PA (AJIN) AJINOMOTO KK.  
 PI Hamura J, Kanayama Y, Nakazawa H, Shimmura T, Sugamura K;  
 PI Takeshita T;  
 DR WPI: 94-325948/41.  
 DR P-PsDB: R60780.  
 PT Immunosuppressant polypeptide - has ability to block  
 PT Interleukin-2 response  
 PS Claim 17; Page 29; 37pp; English.  
 CC Mab capable of binding to the gamma chain of the IL-2 receptor, and  
 CC thus of blocking the IL-2 response, is produced by mouse hybridoma  
 CC line GP-2 (FERM BP-4641). DNA encoding the variable region of  
 CC this Mab was expressed in E. coli, yielding Fv(GP-2) with  
 CC immunosuppressive activity.  
 SQ Sequence 732 BP; 180 A; 204 C; 182 G; 166 T;

Query Match 60.0%; Score 432; DB 1; Length 732;  
 Best Local Similarity 76.7%; Pred. No. 5.2e-106;  
 Matches 557; Conservative 0; Mismatches 160; Indels 9; Gaps 2;

QY 1 ATGCAGCTACCCAGCTCTCAGACATCATGTCTGCATCTCCAGGGGAAAAGTCCACCAG 60  
 DB 7 ATCTCTGCTACCCAGCTCTCAGACATCATGTCTGCATCTCTAGGAGGAGGCTCACCAGT 66  
 QY 61 ACTGAGTGGCCAGCTCAAGTGA-----AGTAACATGACACTGATCAGAGAGATGA 114  
 DB 67 ACCTGAGTGGCCAGCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 126  
 QY 115 AGCAGCTCCCGCAAACTCTGGGTTTATGACACATCCAACTGGCTTCTGGAGTCCAGGT 174  
 DB 127 GGATCTCTCCCGCAAACTCTGGATTTATAGACATCCAACTGGCTTCTGGAGTCCAGGT 186  
 QY 175 CGCTTAGTGGAGTGGTCTGGAACCTTACTCTCTACAGTACGACAGTGAAGCT 234  
 DB 187 CCTTCTAGTGGAGTGGTCTGGAACCTTACTCTCTACAGTACGACAGTGAAGCT 246  
 QY 235 GAAGATGCTGCACCTATTATGTTATGAGGGGAGTGGGATCCATCCAGTTCGGCTCG 294  
 DB 247 GAAGATGCTGCACCTATTATGTTATGAGGGGAGTGGGATCCATCCAGTTCGGCTCG 306  
 QY 295 GGGCAAACTGGAAATTAATAAGTTCTA--CCTCCGATCTGTTAAATCTTGAAGGT 351  
 DB 307 GGGCAAACTGGAGTCTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 366  
 QY 352 AAAGTGTGACGTGACGAGTCTGGGGTTGAGCTTGTCCGAGAGGGGCTTAGTCAAG 411  
 DB 367 AGCAGAGTGAAGTCTGAGAGTCTGGATCTGAGCTGTGAGGCTGAGAGTCTGAGAGTGA 426  
 QY 412 TTGTCTGAAAGCTTCTGACTTCAACATTAATAAGTATTAATATTAATGAGTGAAGAG 471  
 DB 427 TTGTCTGAAAGCTTCTGACTTCAACATTAATAAGTATTAATATTAATGAGTGAAGAG 486

QY 472 AGCCTGAACAGGCGCTGGAAATGATGTGATGATTCATCCTGAGAAATGTAATACTGTA 531  
 DB 487 AGGCATGAGCAAGCGCTGAGTGAAGTGAATTTATTTCTGTAAGTGAAGTGAAGTGA 546  
 QY 532 TATGACCCGAATTTCCAGGGGCAAGGCGAGTATTAAGACAGACATCTCTCCAGCGGCC 591  
 DB 547 TACGATGAGAACTTCAAGGCAAGGCGAGTATTAAGACAGACATCTCTCCAGCGGCC 606  
 QY 592 TACCTTACCTGAGGAGCGCTGACATCTGAGACAGTCCGCTTATTAAGTCTTAT 651  
 DB 607 TACATGACACTTACGACGCTGACATCTGAGAGCTCTGGGTCTATTACTGTCAAGAAC 666  
 QY 652 TACTACTATAGTGTCTACTATGATGATGATGATGATGATGATGATGATGATGATGAT 711  
 DB 667 ACCCGGAGTGGGTCTACTATCTATGATGATGATGATGATGATGATGATGATGATGAT 726  
 QY 712 TCCTCA 717  
 DB 727 TCCTCA 732

## RESULT 6

QY 055180  
 ID 055180 standard; cDNA; 711 BP.  
 AC 055180;  
 DT 21-JUL-1994 (first entry)  
 DE Sequence encoding the single chain anti-erbB2 antibody, Ab no.23.  
 KW Single chain anti-erbB1 antibody; cancer therapy; prevention;  
 OS monoclonal antibody; ss.  
 FH Key  
 FT cds  
 FT MO9400136-A.  
 PD 06-JAN-1994.  
 PF 21-OCT-1992; U08545.  
 PR 30-JUN-1992; US-906555.  
 PA (MOLE) MOLECULAR ONCOLOGY INC.  
 PI Kasprzyk PG, King CR;  
 DR WPI: 94-025878/03.  
 DR P-PsDB: R45442.  
 PT Treatment of malignancies over-expressing ERB-12 - using at least  
 PT 2 monoclonal antibodies which recognise different epitopes on  
 PT gp185  
 PS Example: Fig 7; 37pp; English.  
 CC The source of human erbB-2 protein for the prodn. of antibodies no.  
 CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell  
 CC engineered to express the human erbB-2 protein on its surface (N/  
 CC erbB-2). Abs no. 21 and 23 are directed against the extracellular  
 CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly  
 CC growing tumours were used in a trial of the efficacy of the Abs. In  
 CC animals given a combination of the 2 Abs, tumours completely  
 CC regressed after 11 days.  
 SQ Sequence 711 BP; 175 A; 183 C; 187 G; 165 T;

Query Match 59.7%; Score 429.8; DB 1; Length 711;  
 Best Local Similarity 79.2%; Pred. No. 2e-105;  
 Matches 509; Conservative 1; Mismatches 133; Indels 0; Gaps 0;

QY 2 TGCAGTACCCAGCTCTCCAGCAATCATGTCTGATCTCCAGGGGAAAAGTCCACCATGA 61  
 DB 8 TGCAGTACCCAGCTCTCCAGCAATCATGTCTGATCTCCAGGGGAAAAGTCCACCATGA 67  
 QY 62 CCTGAGTGCACAGCTCAAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGAT 121  
 DB 68 CTTCGAGGGGCCACCCCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 127  
 QY 122 CCCCCAACTCTGGGTTTATGACATCCAACTGGCTTCTGAGATCCAGGTGCTCTCA 181  
 DB 128 CCCCCAACTCTGGGTTTATGACATCCAACTGGCTTCTGAGATCCAGGTGCTCTCA 187  
 QY 182 GTGCAAGTGGTCTTGAAACTCTTACTCTCTACGATCAGCAGATGAGAGCTGAAGATG 241

Db	188	GTGGCGGGGTCTGGGACCTTACTCTCTCACAAGTCACAGAGTGGAGGCTGAAAGATG	24.7
QY	242	CTGCCACTTATTTATTTGTTATTCAGGGGAGTGGGTACCCATTACGCTTCGGGTGGGACAA	30.11
Db	248	CTGCCACTTATTTACTGCCAGAGTGGAGTGGTACCCACACGCTTGGAGGGGGTCCCA	30.77
QY	302	AGTTGGAAATTAAGGTTCTTACCTCCGGATCTGGTAAATCTTTGGAAGGTAAAGGTGTC	36.11
Db	308	AGCTGGAAATTAAGGTTCTTACCTCTGGTTCTGGTAAATCTTTCGAAGGTAAAGGTGTGC	36.75
QY	362	AGCTGCAGCACTTCAGGGGTTGAGTTGTGTCGAGAGGAGGGGCTTGTGTCAGATTGCTCGCA	42.11
Db	368	AGCTGCAGGATTCAGGACCTGAGGTGGTGAAGCTGGAGGTTCAATAAATATCTTCGCA	42.77
QY	422	AAGCTTCTGACTTCACATTTAAAGTATTTATATCCATGGGTGAACAGAGGCTTAAC	48.11
Db	428	AGATTTCTGGTTACTCAATTACTCTGCGCACACAGTAAGTCTGGGTGAACAGGCAATGGA	48.7
QY	482	AGGCGCTGGAAATGGATTGGATGATTCATCTCGAAGTGTAAATCTGTATATGACCGGA	54.11
Db	488	AGAACCTTGATGGATTGGATGACTTATTATCTCTTACATGTGATGATTAACATACACAGA	54.7
QY	542	AATTCACAGGGCAAGGCCAGTTAACACAGACACATCTCTCAACGGCGCTTACCTTAGC	60.11
Db	548	AGTTCAAGGGCAAGGCCACATTTACTGTAGCAAGTGTCTCAGCACGCTTACATGAGC	60.77
QY	602	TCACAGAGCTGACATCTGAGACACTCGCCGCTTTACTGTTGC	64.4
Db	608	TCTCAGCTCTACATCTGAGGACTCTGTGACTCTTATTTACTGTGC	65.0

## RESULT 7

ID Q12405 standard; DNA; 738 BP.

DT 16-SEP-1991 (first entry)

KW Immunoglobulin; antibody; variable region; PE40; immunotoxin

OS Synthetic.

FT	cds	7. .729
FT		442

```
FT misc_rna 7.:.327
```

FT	/label= light chain
EE	328 350

```
/*tag= b
```

FT	misc_rna	370. .729
FT		44.3

```

FT      /label= heavy chain
EX      rct7463111-1

```

PD 28-MAY-1991. 122066  
 TT 11-TAN-1000. 122066

PR 11-JAN-1990; US-463111.

PI Pastan I;  
RFB WBI: 01-300877/337

DR P-PSDB; R12798.

by fusing genes with DNA encoding c

PS Disclosure; Fig 3; 26pp; English.

CC chains of murine Mab OVB3, joined to

CC linked to the 5' end of the coding

CC	clone can be used to kill specific	
60	sequence	738 bp
		191
		103

Query	6	GCTGACCCAGTCTCCAGCAATCATGTCTGCACTCTCCAGGGGAAAAAGTCCACCATGACTG	65
Db	15	GCTACCCAGTCTCCAGCAATCATGTCTGCACTCTCCAGGGGAAAAAGTCCACCATGACTG	74
QY	66	CAGTGCAGCTCAAGTGTAACTAATGACATGCTATACAGAGAAAGTCAACACCTCTCCC	125
QY	75	CAGTGCAGCTCAAGTGTAACTAATGACATGCTATACAGAGAAAGTCAACACCTCTCCC	134

Query Match 56.2%; Score 404.4; DB 1; length 738;  
 Best Local Similarity 73.7%; Pred. 0.1, ie-98;  
 Matches 529; Conservative 0; Mismatches 186; Indels 3; Gaps 1.

OY	6	GCTACCCAGAGCTCCACACATCAATGTCGATGCCGAGGGGAAAAAGGTCAACATGACCTG	65
Db	15	GCTACCCAGAGCTCCACACATCAATGTCGATGCCGAGGGGAAAAAGGTCAACATGACCTG	74
OY	66	CAGTCCAGGCTCAAGTGTAAGTACATGACACTGGTATACACAGAAATGCAAGCACTCC	125
Db	75	CAGTCCAGGCTCAAGTGTAAGTACATGACACTGGTATACACAGAAATGCAAGCACTCC	134
OY	126	CAAACTCTGGGTTTATGACACATCCAAACTGGCTTGGAGTCCCAAGTCGCTTCACTG	185
Db	135	CAAACTCTGGGTTTATGACACATCCAAACTGGCTTGGAGTCCCAAGTCGCTTCACTG	194
OY	186	CAGTGGGCTCGGAAACCTTACTCTCAGAGTACAGACATGAGAGCTGAATGCTGC	245
Db	195	CAGTGGGCTCGGAAACCTTACTCTCAGAGTACAGACATGAGAGCTGAATGCTGC	254
OY	246	CACCTATATGTTATCAGAGGGATGGGTACCCATTCAGCTGGCGCTCGGGACAAAGTT	305
Db	255	CACCTATATGTTATCAGAGGGATGGGTACCCATTCAGCTGGCGCTCGGGACAAAGTT	314
OY	306	GGAATATAAA---GGTCTACCTCCGAGTCTGTAAATCTCTTGAAGSTAAAGGTGTGCA	362
Db	315	GGAGCTCAAAAGAGGGTAATCCTAGAGATGCGCTCGGAATCCAAATCGACAGATGCGCA	374
OY	363	GCTGCAGCAGCTCTGGGGTTGAGCTGTGCCAGAGAGGGGCTTAGTCAAGTTGCTGCAA	422
Db	375	GCTGGTGGAGCTGTGGGGAGGCTTAGTGCAGCTCGAGAGGGTCCCGCAACTCTCTGTGC	434
OY	423	AGCTTCAGCTCAACATTTAAAGCTATTATATCCACTGGTGAAGAGAGAGCCTGAAACA	482
Db	435	AGCCTCTGATTCACCTTCAGTAAGCTTTGGAATCCACTGGTTCGTGAGGCTCCAGAGAA	494
OY	483	GGGCTCGGAATGGATGGATGGATATCCGGAATGGTATATCTGTATATGACCCGAA	542
Db	495	GGGCTCGGAATGGATGGATGGATATCCGGAATGGTATATCTGTATATGACCCGAA	554
OY	543	ATTCCAGGGCAAGCCAGATATACAGACAGACATCTCCACAGCGGCTACCTTACGCT	602
Db	555	AGTGAAGGGCCGATTCACACATCTCCAGAGAACATCCCAAGAACCCGTGTTCCGCAAT	614
OY	603	CAGCAGCTGCATCTAGGAGCACTGCCGTCTATTACTGTGCTTCTTATTAATCTACTATG	666
Db	615	GACCAGCTTAAGCTCTGACGACAGCGGCATGATTAATCTGTGCAAGATCCCTTACTAGG	674
OY	663	TGCTTACATAGCTATGTAATCTACTGGGGCAAGAAAGCTGGTCAAGCCTTCCTCATATA	720
Db	675	CTACTCTATCTTATGAGCACTCTGGGGCAAGGAACCACTCTCAAGTCTCTCTTAAAA	732

## RESULT

ID T91615 standard; cDNA to mRNA; 720 BP

DT 22-DEC-1997 (first entry)

KW Single chain Fv; scFv; antibody; *E. coli*; anti-T3 antibody;

05 Synthetic.

PD 26-AUG-1997.

PR 15-FEB-1996; JP-027622.

DR WPI; 97-474306/44.

PT Producing single chain Fv antibody in *Escherichia coli* - by

PT co-expression with a chaperonin as a soluble fraction

PS Claim 4; Page 6-7; 9pp; Japanese.  
CC The sequences given in T91614-15 encode single chain Fv (scFv)  
CC antibodies which are produced in E. coli. The scFv's are derived  
CC from an anti-T3 antibody or an anti-gp130 antibody and are produced  
CC either by: (1) expression as an inclusion body, followed by folding  
CC (i.e. by denaturation and solubilisation) or (2) expression as a  
CC soluble fraction by co-expression with chaparotin. Using the methods,  
CC scFv can be produced in E. coli, either in a soluble fraction or in  
CC inclusion bodies. Like the parent double chain antibodies, the scFv  
CC polypeptides are useful in immunological diagnosis or for separation  
CC and purification, but they lessen the side effects caused by antibody  
CC constant regions. The availability of scFv polypeptides contributes  
CC to fundamental research and development of diagnostic and therapeutic  
CC drugs.  
SO Sequence 720 BP; 169 A; 193 C; 190 G; 168 T;  
  
Query Match 53.3%; Score 383.8; DB 1; Length 720;  
Best Local Similarity 73.0%; Pred. No. 3.2e-93;  
Matches 522; Conservative 0; Mismatches 187; Indels 6; Gaps 2;  
  
OY 6 GGTGACCCAGTCTCCAGCAATCATGCTCATCTCCAGGGGAAAGTCCACATGACCTG 65  
DB 9 GCTCACCCAGTCTCCAGCAATCATGCTCATCTCCAGGGGAAAGTCCACATGACCTG 68  
OY 66 CAGTCCAGCTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 125  
DB 69 CAATGTCAGCTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 128  
OY 126 CAAATCTGGGGTTTATGACATCCAAATGCTTCTGATGATGATGATGATGATGATG 185  
DB 129 CAAACCTGGGGTTTATGACATCCAAATGCTTCTGATGATGATGATGATGATGATG 188  
OY 186 CAGTGGGTGGAATCTTACTCTCTCAGATCAGATCAGATCAGATCAGATCAGATCAG 245  
DB 189 CAGTGGGTGGAATCTTACTCTCTCAGATCAGATCAGATCAGATCAGATCAGATCAG 248  
OY 246 CAGTGGGTGGAATCTTACTCTCTCAGATCAGATCAGATCAGATCAGATCAGATCAG 305  
DB 249 CAGTGGGTGGAATCTTACTCTCTCAGATCAGATCAGATCAGATCAGATCAGATCAG 308  
OY 306 GGAATTAAGTGTCTACTCTCGGATCTGGTAAATCTTCTGAAAGTAA--GGTGTGCA 362  
DB 309 CGAGCTGGGTGGGCTGCTCGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 368  
OY 363 GCTGACAGAGTCTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 422  
DB 369 ATTGACAGAGTCTGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 428  
OY 423 AGCTTCTGACTTCAACATTAAAGTATTAATTCACAGTGGTGAAGCAGAGGCTGAACA 482  
DB 429 GCTTACAGAGTACATCTACGACTACACAAATGAGCTGGGTGAAGCAGAGGCTGAACA 488  
OY 483 GGGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 542  
DB 489 GAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 548  
OY 543 ATTGACAGAGGCTGAGATTAACAGACAGACATCTCCAGAGGGGCTTACCTTACAGT 602  
DB 549 GTTCAAGGACAAAGGCTATTAATTAAGTCTTACAGACAGGCTTACATGAGCT 608  
OY 603 CAGAGCTGACATCTGAGGACAGTCCGCTTATTAATGATGATGATGATGATGATGATG 662  
DB 609 CCGAGAGCTGACATCTGAGGACAGTCCGCTTATTAATGATGATGATGATGATGATG 665  
OY 663 TGTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 717  
DB 666 GTACGGCACTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 720

AC T91614;  
DT 22-DEC-1997 (first entry)  
DE cDNA encoding an anti-T3 antibody derived scFv.  
KW Single chain Fv; scFv; antibody; E. coli; anti-T3 antibody;  
KW anti-gp130 antibody; inclusion body; chaparotin; ds.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT cds 1..735  
FT /tag- a  
FT /trans\_except- pos:160..173, aa:TYT  
FT /trans\_except- pos:385..387, aa:GLU  
FT /note- "No stop codon given"  
FT  
FN J09220092-A.  
PD 26-AUG-1997.  
PF 15-FEB-1996; 027622.  
PR 15-FEB-1996; JP-027622.  
RA (TOYO) TOSOH CORP.  
DR WP1: 97-474306/44.  
DR P-PSDB; W25783.  
PT Producing single chain Fv antibody in Escherichia coli - by  
PT co-expression in an inclusion body, followed by protein folding or by  
PT co-expression with a chaparotin as a soluble fraction  
PS Claim 3; Page 5-6; 9pp; Japanese.  
CC The sequences given in T91614-15 encode single chain Fv (scFv)  
CC antibodies which are produced in E. coli. The scFv's are derived  
CC from an anti-T3 antibody or an anti-gp130 antibody and are produced  
CC either by: (1) expression as an inclusion body, followed by folding  
CC (i.e. by denaturation and solubilisation) or (2) expression as a  
CC soluble fraction by co-expression with chaparotin. Using the methods,  
CC scFv can be produced in E. coli, either in a soluble fraction or in  
CC inclusion bodies. Like the parent double chain antibodies, the scFv  
CC polypeptides are useful in immunological diagnosis or for separation  
CC and purification, but they lessen the side effects caused by antibody  
CC constant regions. The availability of scFv polypeptides contributes  
CC to fundamental research and development of diagnostic and therapeutic  
CC drugs.  
SO Sequence 735 BP; 165 A; 192 C; 197 G; 181 T;  
  
Query Match 45.3%; Score 326.4; DB 1; Length 735;  
Best Local Similarity 67.6%; Pred. No. 5.5e-78;  
Matches 491; Conservative 0; Mismatches 226; Indels 9; Gaps 2;  
  
OY 1 ATGACGCTGACCCAGTCTCCAGCAATCATGCTCATCTCCAGGGGAAAGTCCACATG 60  
DB 10 ATTGACGCTGACCCAGTCTCCAGCAATCATGCTCATCTCCAGGGGAAAGTCCACATG 69  
OY 61 ACCTGCACTGCTCAAGTGTAA--TAAATGCACTGCTATCAGCAGAAATCA 114  
DB 70 ACCTGCACTGCTCAAGTGTAA--TAAATGCACTGCTATCAGCAGAAATCA 129  
OY 115 AGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 174  
DB 130 GATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189  
OY 175 CGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 234  
DB 190 CGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 249  
OY 225 GAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294  
DB 250 GAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309  
OY 295 GGGAAAGTGGGAAATTAAGTGTACCTCCGATGCTGCTGCTGCTGCTGCTGCTGCTG 354  
DB 310 GGGAAAGTGGGAAATTAAGTGTACCTCCGATGCTGCTGCTGCTGCTGCTGCTGCTG 369  
OY 355 --GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411  
DB 370 TCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429  
OY 412 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471



Db 430 CTCCTGTGAAGCCTTGATTCACCTTTCAGTAGTTATTACATGTCCTTGGTGGCCAG 489  
QY 472 AGGCTGAAACAGGCGCTGGAATGATGATGATTCATCCGAGATGTAATTAATGTA 531  
Db 490 ACTCCAGAGAGAGAGGCTGTGGTGGCAGCCATTAATAGTATGCTACCACTTAC 549  
QY 532 TATGACCGAAATTCAGAGGCAAGGCCAGTATACAGACATCCCTCCAAAGGGGCC 591  
Db 550 TATTCAGACACTGTGAAGGCGGATTCACCATCTCCAGACAAATGCAAGAACCTTG 609  
QY 592 TACCTTCAGCTGACGACCTGACATCTGAGACATGCGCTTATTAATGCTTCTTAT 651  
Db 610 TACCTTCGAATGAGAGCTGTGAGTCTGAGACACAGCCCTTATTAATGCTGACCCG 659  
QY 652 TACTACTATAGTCTTACTATGCTATGATGCTACTGAGGCTCAAGAACTGCTACCCGTC 711  
Db 670 GTCCTCTATTAATGCTATGCTATGCTATGCTACTGAGGCGCAAGGAGGCTCACCGTC 729  
QY 712 TCCTCA 717  
Db 730 TCCTCA 735

## RESULT 10

Q51540  
ID 051540 standard; DNA; 720 BP.  
AC Q51540;  
DE 23-MAY-1994 (first entry)  
DI Coding sequence of polypeptide with affinity for fluorescein.  
KW Monoclonal antibody; Mab; affinity; binding; antigen; diagnostics;  
KW therapy; imaging; purification; biosensors; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 619..621  
FT FT /transl\_except- AAT encodes Isoleucine.  
FT FT /tag- a  
FT FT misc\_difference 658..660  
FT FT /transl\_except- GAA encodes Glycine.  
FT FT /tag- b  
FT FT /transl\_except- GAA encodes Glycine.  
PN US5260203-A.  
PD 09-NOV-1993.  
PF 02-SEP-1986; 902971.  
PR 02-SEP-1986; US-902971.  
PR 02-SEP-1987; US-092110.  
PR 19-JAN-1989; US-299617.  
PR 25-APR-1990; US-512910.  
PA (ENZO-) ENZO INC.  
PI Bird RE, Hardman K, Ladner RC;  
DR WPI; 93-367875/46.  
DR P-PSDB; R43679.  
PT Single chain polypeptide for binding antigen - comprising light  
PT and heavy chain antigen binding portions linked by peptide linker  
PS Example 9, Figure 39; 78pp; English.  
CC The single chain polypeptide is derived from the mature light and  
CC heavy chains of a monoclonal antibody (Mab) and has affinity  
CC for a given antigen (Fluorescein). It comprises a first  
CC polypeptide comprising the antigen binding portion of of the light  
CC chain variable region of an antibody and a second polypeptide  
CC comprising the antigen binding portion of the heavy chain variable  
CC region of an antibody and at least one peptide linker linking the  
CC first and second polypeptide chains. The resulting single chain  
CC polypeptide can be used in diagnostics, therapy  
CC (in vivo and in vitro), imaging, purifications and biosensors.  
CC This particular single chain binding molecule was designated  
CC 18-2-3/TRY202' and contains one linker peptide.  
SQ Sequence 720 BP; 180 A; 185 C; 175 G; 180 T;

## Query Match

44.7%; Score 322; DB 1; Length 720;

Best Local Similarity 68.7%; Pred. No. 8e-77;  
Matches 493; Conservative 0; Mismatches 210; Indels 15; Gaps 3;

QY 6 GCTGACCACTCTCCAGCAATCATGTCGATCTCCAGGGAAGGTCACATGACCTG 65

Db 12 GCTACCCAGTCTCCAGCAATCATGTCGATCTCCAGGGAAGGTCACATGACCTG 71  
QY 66 CAGTGCACCTCAAGTGA-----AGTAACATGACCTGATATCAAGACAAAGCAAC 119  
Db 72 CAGGCGACCTCAAGTGAAGTTCAGAGTACTTGCAGCTGTACAGCAAGAACTCAGGTG 131  
QY 120 CTCGCCCAACTCTGGGTTTATGACACATCCAAATGCTGCTTGAGATGCCAGTGCCTT 179  
Db 132 CTCGCCCAACTCTGGGTTTATGACACATCCAAATGCTGCTTGAGATGCCAGTGCCTT 191  
QY 180 CAGTGCACCTCAAGTGAAGTTCAGAGTACTTGCAGCAATGACAGATGAGAGGCTGA 239  
Db 192 CAGTGCACCTCAAGTGAAGTTCAGAGTACTTGCAGCAATGACAGATGAGAGGCTGA 251  
QY 240 TCGTGCACCTCAAGTGAAGTTCAGAGTACTTGCAGCAATGACAGATGAGAGGCTGA 299  
Db 252 TCGTGCACCTCAAGTGAAGTTCAGAGTACTTGCAGCAATGACAGATGAGAGGCTGA 311  
QY 300 AAAGTTGGAATTAAGGTTTACTCTCGGATCTGGTAAATCTTCTGAAGTAAAGTGT 359  
Db 312 CAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 365  
QY 360 GCAAGTGCAGCACTCTGGGTTGAGCTTCCGAGAGAGGCGCTTATGATCAAGTGTCTG 419  
Db 366 TCAGCTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 425  
QY 420 CAAAGCTCTGACTTCAACATTAAGACATTAATTAATCACTGAGGTAAGACAGGCTGA 479  
Db 426 CACCGCTCTGGTGTTCATTAACCACTATGATGCTGACCTGGTTCGCCAGCTTCAG 485  
QY 480 ACAGGCGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539  
Db 486 AAAGGCTGAGATGCTGGAGTATATATGAGCTG---TGAAACACAATTAATATTC 542  
QY 540 GAAATTCAGGCGCAAGGCCAGATTAACACACACATCTCCAAAGCGGCTTACCTTA 599  
Db 543 AGCTTCATGTCAGACATGACATGACATGACATGACATGACATGACATGACATGAC 602  
QY 600 GCTCAGCAGCTGACATCTGAGGACATGCGGCTTACTGCTGCTTATTACTACTA 659  
Db 603 AATGACAGTCTGCAAAATGATGACACAGCCTATTAATCTGCTGCTGCTGCTGCTG 662  
QY 660 TACTGCTTACTATGCTATGCTACTGAGGCTCAAGAACTGCGTCAAGCTCTCTCA 717  
Db 663 AATCTTTACTATGCTATGAGATGATGAGGCTCAAGAACTGAGTCAAGCTCTCTCA 720

## RESULT 11

T13739  
ID T13739 standard; DNA; 720 BP.  
AC T13739;  
DE 11-OCT-1996 (first entry)  
DI Single chain binding molecule 18-2-3/TRY202' DNA.  
KW Antibody engineering; single polypeptide chain binding molecule;  
KW single chain antibody; SCA; heavy chain; light chain;  
KW monoclonal antibody; Mab; immunofluorescence purification;  
OS Chimeric murine sp.  
OS Chimeric synthetic.  
FH Key Location/Qualifiers  
FT cds 619..621  
FT FT /tag- a  
FT FT /transl\_except- (619..621, aa:11e)  
FT FT /tag- b  
FT FT /transl\_except- (658..660, aa:gly)  
FT FT /tag- b  
FT FT /transl\_except- (658..660, aa:gly)  
PN US534621-A.  
PD 09-JUL-1996.  
PF 02-SEP-1986; 902971.  
PR 02-SEP-1986; US-902971.  
PR 02-SEP-1987; US-092110.  
PR 19-JAN-1989; US-299617.

PR 25-APR-1990: US-512910.  
 PR 01-APR-1993: US-040440.  
 PR 06-JUN-1995: US-468992.  
 PA (ENZO-) ENZON LABS INC.  
 PI Bird RE, Hardman K, Ladner RC;  
 DR WPI: 96-33309/73.  
 DR P-PSDB: R99649.  
 PT Immunopurificn. using single binding chain molecule including  
 PT antigen-binding parts of antibody light and heavy chain variable  
 PT regions connected by a linker - is smaller, stabler and less  
 PT expensive than complete antibodies  
 PS Example 9: Fig 39: 78pp: English.  
 CC A DNA construct (713739) codes for single chain binding molecule  
 CC 18-2-3/TRY202 (R99649), in which VL and VH regions of anti-  
 CC fluorescein monoclonal antibody 18-2-3 are joined by a peptide  
 CC linker composed primarily of alternating Gly and Ser residues, with  
 CC Glu and Lys residues to enhance solubility. The construct was  
 CC expressed in E.coli. Soluble, folded chain antibody (SCA) was  
 CC obtd. which was capable of exhibiting a biological binding  
 CC activity equivalent in specificity and affinity to that of a  
 CC monoclonal antibody. The SCA has the advantages of smaller size,  
 CC greater stability and reduced cost.  
 SQ Sequence 720 BP; 180 A; 184 C; 175 G; 181 T;

Query Match 44.7%; Score 322; DB 1; Length 720;  
 Best Local Similarity 68.7%; Pred. No. 8e-77;  
 Matches 493; Conservative 0; Mismatches 210; Indels 15; Gaps 3;

OY 6 GGTGACCAAGTCTCCAGCAATCATGCTGATCTCCAGGGGAAAGGTCACATGACCTG 65  
 DB 12 GCTCACCCAGTCTCCAGCAATCATGCTGATCTCCAGGGGAAAGGTCACATGACCTG 71  
 OY 66 CAGTGCACAGTCTCAAGTGTGTA-----AGTAACATGACAGCTGATACAGAGTCAAGCAC 119  
 DB 72 CAGGCGCAGCTCAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 131  
 OY 120 CTCCTCCCAAACTGGGTTTATACACATCCAACTGGCTTCTGAGTCCAGCTGCTT 179  
 DB 132 CTCCTCCCAAACTGGGTTTATACACATCCAACTGGCTTCTGAGTCCAGCTGCTT 191  
 OY 180 CAGTGCAGTGGGTGGAAGTCTTACTCTCCAGAGTCAAGAGATGAGGCTGAAGA 239  
 DB 192 CAGTGCAGTGGGTGGAAGTCTTACTCTCCAGAGTCAAGAGTGGAGGCTGAAGA 251  
 OY 240 TGTGCGCACTTATTTATGTTATACAGGGGAGTGGTACCCATTCAGCTTGGCTGGGAGC 299  
 DB 252 TGTGCGCACTTATTTATGTTATACAGGGGAGTGGTACCCATTCAGCTTGGCTGGGAGC 311  
 OY 300 AAGTTGGAATTAAGGTTCTACTCCGATCTGTAATCTTCTGAAGGTAAAGGTGT 359  
 DB 312 CAGGCTTGAAGTGAAGGTAAATCTTCTGTTCTGG-----TTCCGAATCTTAATCTTAC 365  
 OY 360 GCAGCTGACAGCACTGGGGTTGAGCTTGCCAGAGAGGGGCTTATGTCAGTTCGTC 419  
 DB 366 TCGCTGAAAGAGTCAAGAGCTCTCTGAGGCGCCCTACAGAGCCTGTCTCACTTGC 425  
 OY 420 CAAAGCTTGTGACTTCAATTAAGACTATTTATATCACTGGGTGAAGCAGAGGCTGTA 479  
 DB 426 CACTGCTCTGGGTTTCAATTAAGCACTATGTTGACACTGGGTTCGCCAGCTCCAGG 485  
 OY 480 ACAGGCGCTGGAAATGATGATGATGATCTATCTGAGAAATGTAATCTGTATATGACCC 539  
 DB 486 AAGGCGCTGAGTGGGTGAGTAATATGAGCTGG---TGAAGAACAAATATTAATTC 542  
 OY 540 GAAATTCAGAGGAGCCAGTATACAGAGACACATCTCCAGAGCGGCTCACTTCA 559  
 DB 543 AGCTTCATGTCTCAGACTGACATGACCAAGCAATTCCTCAAGAGCCAAAGTTTCTTAA 602  
 OY 600 GCTCAGCAGCTGACATCTGAGGAGCACTGGCTTATCTACTGCTTCTTATCTACTACTA 659  
 DB 603 TATGAAACAGTCTGCAAAATGATACACAGCATACTACTGTGCAAGAGCACTGGAAG 662

OY 660 TAGTGTACTATGATATGATCTACTGAGGCTCAAGAGAACCTGGTCCGCTCTCTCA 717  
 DB 663 AATCTTACTATGCTATGACTATTTGGGTCAAGAGAACCTGATCAGCTCTCTTAA 720  
 RESULT 12  
 ID T36463 standard; DNA: 720 BP.  
 AC T36463;  
 DT 13-NOV-1996 (first entry)  
 DE 18-2-3/TRY202' single chain binding protein-encoding DNA.  
 KW Antibody engineering; monoclonal antibody; Mab.  
 KW single chain antibody; immunosassay; fluorescein; 18-2-3/TRY202';  
 OS Chimeric Mus musculus;  
 FH Chimeric synthetic.  
 FT Location/Qualifiers  
 FT cds 558..660  
 FT /tag\_a  
 FT /transl\_except- (658..660, aa:Gly)

PN US518889-A.  
 PD 21-MAY-1996.  
 PF 02-SEP-1986: 902971.  
 PR 02-SEP-1986: US-902971.  
 PR 02-SEP-1987: US-092110.  
 PR 19-JAN-1989: US-299617.  
 PR 25-APR-1990: US-512910.  
 PR 01-APR-1993: US-040440.  
 PR 06-JUN-1995: US-468988.  
 PA (ENZO-) ENZON LABS INC.  
 PI Bird RE, Hardman K, Ladner RC;  
 DR WPI: 96-259060/26.  
 DR P-PSDB: W02191.  
 PT Immunossay using single chain antigen binding mol. - as replacement  
 PT for labelled or immobilised antibody, are less immunogenic, easier  
 PT to engineer, more stable and less expensive  
 PS Example 9: Fig 39A-B: 78pp: English.  
 CC Computer-designed construct 18-2-3/TRY202' (W02191), encoded by  
 CC the DNA sequence given in T36463, comprises the variable regions  
 CC of the light and heavy chains of anti-fluorescein monoclonal  
 CC antibody (Mab) 18-2-3 linked by a peptide designed to fit into a  
 CC groove on the backside of the variable domain structure. The  
 CC 18-2-3 VH and VL regions were obtd. by PCR amplification of  
 CC hybridoma RNA, and the DNA construct was inserted into vector  
 CC pGX3703 and introduced into E. coli. 18-2-3/TRY202' was  
 CC expressed as a single chain molecule. It exhibited biological  
 CC binding activity equivalent in specificity and affinity to  
 CC that of the original Mab.  
 SQ Sequence 720 BP; 180 A; 184 C; 175 G; 181 T;

Query Match 44.7%; Score 322; DB 1; Length 720;  
 Best Local Similarity 68.7%; Pred. No. 8e-77;  
 Matches 493; Conservative 0; Mismatches 210; Indels 15; Gaps 3;

OY 6 GGTGACCAAGTCTCCAGCAATCATGCTGATCTCCAGGGGAAAGGTCACATGACCTG 65  
 DB 12 GCTCACCCAGTCTCCAGCAATCATGCTGATCTCCAGGGGAAAGGTCACATGACCTG 71  
 OY 66 CAGTGCACAGTCTCAAGTGTGTA-----AGTAACATGACAGCTGATACAGAGTCAAGCAC 119  
 DB 72 CAGGCGCAGCTCAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 131  
 OY 120 CTCCTCCCAAACTGGGTTTATACACATCCAACTGGCTTCTGAGTCCAGCTGCTT 179  
 DB 132 CTCCTCCCAAACTGGGTTTATACACATCCAACTGGCTTCTGAGTCCAGCTGCTT 191  
 OY 180 CAGTGCAGTGGGTGGAAGTCTTACTCTCTCAGATCAGATCAGAGTGAAGGCTGAAGA 239  
 DB 192 CAGTGCAGTGGGTGGAAGTCTTACTCTCTCAAAATCAGAGTGAAGGCTGAAGA 251  
 OY 240 TGTGCGCACTTATTTATGTTATACAGGGGAGTGGTACCATTCACAGTTGGCTGGGAGC 299











```

Oy      122  VOLQSSGEVYVPKSGSMKISCTSGSYFTGHTMMNVKSHGNLEIGLNPNDGTNYN- 181
         |||:||||:||||| ||||| |||||:|||||:|||||:|||||:||||| | : ||
Db       2  VOLQSSGEELVKPKGSMTIKSCASGISFTGTMMNVKSHGSKNNLEWGLNPDKVSITYN 61
         |||:||||:||||| ||||| |||||:|||||:|||||:|||||:||||| | : ||

Oy      182  QKFNKATFTVDKSSSTAYMELLSTSEDSAVYYCARR---VTDMYFEDVWGAGTFTVVS 237
         |||:||||:||||| ||||| |||||:|||||:|||||:|||||:||||| | : ||
         :||||:||||| ||||| |||||:|||||:|||||:|||||:||||| | : ||
Db       62  QKFNKATFTVDKSSSTAYMELLSTSEDSAVYYCARSGYGSDMYFDVWGAGITTVVS 121
         |||:||||:||||| ||||| |||||:|||||:|||||:|||||:||||| | : ||

RESULT   3
I37267
Ig heavy chain V region (129) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 16-Aug-1996
C:Accession: I37267
R.Ruff-Jamison, S.: Campos-Gonzalez, R.; Glenney Jr., J.R.
J. Biol. Chem. 266, 6607-6613, 1991
A>Title: Heavy and light chain variable region sequences and antibody properties of anti-
A:Reference number: A38740; MUID:91177923
A:Accession: I37267
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-128 <RUF>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:19-102/Domain: immunoglobulin homology <IMM>
```

Query Match	39.3%	Score 498;	DB 2;	Length 128;
Best Local Similarity	77.2%;	Pred. No. 5.2e-30;		
Matches 95; Conservative	10;	Mismatches 12;	Indels 6;	Gaps 1.

QY	121	GVOLDSSEPELVKPPGSKKISCKTSGVSTFCHTNAMWVQKSHGKMLWEIGLIPVNCSTY	180
QY	122	GVOLDSSEPELVKPPGSKKISCKTSGVSTFCHTNAMWVQKSHGKMLWEIGLIPVNCSTY	181
Db	5	GVOLDSSEPELVKPPASVYKISCKTSGVSTFETVYTHMWVQKSHGKMLEWIGGINPNNGSTY	64
QY	181	NQKFGKATFTVNDKSSSTAYMELLSTNSEDASVYYCAR-----YADWYEDVWAGSTV	234
Db	65	NQKFGKATFTVNDKSSSTAYMELSLNSEDASVYYCARGLITTVANASYTFDYGQSTTL	124
QY	235	TVS 237	
Db	125	TVS 127	

```

RESULT      4
MHMS4E
1g heavy chain V region (MOPC 104E) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Nov-1979 #sequence_revision 14-Nov-1983 #text_change 16-Aug-1996
C:Accession: A02039
R:Kerry, M.R.; Fuhrman, J.S.; Schilling, J.W.; Rogers, J.; Sibley, C.H.; Hood, L.E.
Biochemistry 21, 5415-5444, 1982
A:Title: Complete amino acid sequence of a mouse mu chain: homology among heavy chain CC
A:Reference number: A02039; MUID:83075344
A:Accession: A02039
A:Molecule type: protein
A:Residues: 1-117 <KEH>
C:Comment: The glycosylated asparagine residue does not have the usual N-X-S/T context a
C:Comment: This protein binds dextran.
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: glycoprotein; heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <TMM>
F:22-96/Disulfide bonds: #status predicted
F:55/Binding site: carbonylrate (asn) (covalent) #status atypical

```

```

Query Match      39.28; SCORE 496.5; DB 1; Length 117;
Best Local Similarity 81.0%; Pred. No. 66-30; 12; Indels 1; Gaps 1;
Matches 94; Conservative
07 122 VOLOSGPEVYVPGGSGMRISCKTSGYSPFGHMNMWVKOSHGNKLEWIGLINPYNGDTNN 181
    |||||:|||||:|:-||| |||:-:-: | |||||:||||| ||| || |:-||

```

Db 2 VQLQSSPELVKRGASVKMSKASGTYTFTDYYMKVVKQSHKSLSEJLGGDINPNNGGISYN 6

Qy 182 QKFKGKATFTVDKSSSTAYMELLSTLSEDSAYVYCARRYTDWYFDVWAGAGTYTVS 237

Db 62 QKFKGKATLTVDKSSSTAYMQLNLSLSEDSAYVYCARDY-DWYFDVWAGAGTYTVS 116

၆

Ig heavy chain V region (Py69) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Oct-1991 #sequence\_revision 18-Oct-1991 #text\_change 16-Aug-1996  
C:Accession: C37267  
R:Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.  
J. Biol. Chem. 265, 6607-6613, 1991  
A:Title: Heavy and light chain variable region sequences and antibody properties of a  
A:Reference number: A38740; MUID: 91j77923  
A:Accession: C37267  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-128 <RUF>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

Query Match	39.1%;	Score 495;	DB 2;	Length 128;
Best Local Similarity	-76.4%;	Pred. NO. 8.5e-30;		
Matches 94;	Conservative 10;	Mismatches 13;	Indels 6;	Gaps 1

OY	121	GVOLQESPEYVVRKSGSKRISCKTSGSFTGHTANWVKQSHGKLTLEWIGLIPNPGCTNY	180
OY	121		180
Db	5	GVOLQOSPEYVVRKSGASKRISCKTSGYTFTEYTHWVKQSHGKLTLEWIGRNPNSGSGTY	64
OY	181	NQKFKGKATFTYDRSSSTAYVWELLSTLSEDSAVYYCAR-----VMDYFEDVWAGCTNY	234
Db	65	NQKFKGKATFTYDRSSSTAYVWELLSTLSEDSAVYYCARRGKGYNTSYTFDYWGQGTTL	124
OY	235	TVS	237
Db	125	TVS	127

```

RESULT      6
PC4405
anti-AFP antibody F3 light chain variable region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 10-Nov-1997 #sequence_revision 23-Jan-1998 #text_change 26-Feb-1998
C:Accession: PC4405
R: Deng, J.B.; Han, H.; Su, C.Z.; Chen, C.Q.
Chinese Biochem. J. 12, 648-653, 1996
A:Title: Generation of a phage display library of the immunoglobulin repertoire from
A:Reference number: PC4405
A:Accession: PC4405
A:Molecule type: mRNA
A:Residues: 1-107 <DEN>
A:Experimental source: spleen cell
C:Superfamily: immunoglobulin V region: immunoglobulin homology

```

Query Match	38.7%;	Score 490;	DB 2;	Length 107;
Best Local Similarity	84.0%;	Pred. No. 1.6e-29;		
Matches	89;	Conservative	9;	Mismatches 8;
			Indels	0;
			Gaps	0

[illegible]

## RESULT 7



E45722  
anti-glycoprotein H monoclonal antibody heavy-chain variable domain (Mab 5) - mouse (fra

C:Species: Mus musculus (house mouse)  
C:Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-May-1997  
C:Accession: E45722  
R:Simpson, J.A.; Chow, J.C.; Baker, J.; Avdalovic, N.; Yuan, S.; Au, D.; Co, M.S.; Vasqu  
J. Virol. 67, 489-496, 1993  
A>Title: Neutralizing monoclonal antibodies that distinguish three antigenic sites on hu  
A:Reference number: A45722; MUID:93100833  
A:Accession: E45722  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-120 <STM>  
C:Note: sequence extracted from NCBI backbone (NCBIP:120593)  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: glycoprotein

Query Match 38.7%; Score 490; DB 2; Length 120;  
Best Local Similarity 80.5%; Pred. No. 1.8e-29;  
Matches 95; Conservative 8; Mismatches 13; Indels 2; Gaps 2;

OY 122 VOLOSGPEVVPKPGSMKISCKTSGSFTGHMMNVKOSHGNLEMGIL-INPYNGDTNY 180  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|  
DB 2 VOLQOSGEPLVKPKGASMKISKASYSTGTGMNVKOSHGNLEMGILINPNIGTSY 61

OY 181 NQFKGKATITVDKSSSTAYMELLSLTSEDSAAVYICARR-VTDWTFDVGAGTTIVTS 237  
|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:  
DB 62 NQFKGKATITVDKSSSNAYMELLSLTSDASDAVYYICTRGFRDYMDFMQGISTVTS 119

RESULT 8  
MHMSJS  
Ig heavy chain V region (J558) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 16-Aug-1996  
C:Accession: A26242  
R:Schilling, J.; Clevinger, B.; Davie, J.M.; Hood, L.  
Nature 283, 35-40, 1980  
A>Title: Amino acid sequence of homogeneous antibodies to dextran and DNA rearrangements  
A:Reference number: A26242; MUID:80078170  
A:Accession: A26242  
A:Molecule type: protein  
A:Residues: 1-117 <SCH>  
A:Note: The sequences of 10 hybridoma proteins that also bind dextran differ from that s  
C:Comment: This protein binds dextran.  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; hybridoma; immunoglobulin  
F:15-98/Domain: immunoglobulin homology <IMM>  
F:22-96/Disulfide bonds: #status predicted

Query Match 38.5%; Score 487.5; DB 1; Length 117;  
Best Local Similarity 80.2%; Pred. No. 2.7e-29;  
Matches 93; Conservative 9; Mismatches 13; Indels 1; Gaps 1;

OY 122 VOLOSGPEVVPKPGSMKISCKTSGSFTGHMMNVKOSHGNLEMGILINPYNGDTMYN 181  
DB 2 VOLQOSGEPLVKPKGASKASGTTFTDYMKWKVKSHGKSLEMIGLINPNNGSTIN 61

OY 182 QKFEGKATITVDKSSSTAYMELLSLTSEDSAAVYICARRTDMVFDPWGAGTTIVTS 237  
|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:  
DB 62 QKFEGKATITVDKSSSTAYMQLNSLTSEDSAAVYICARD-RYYTFDWGAGTTIVTS 116

RESULT 9  
PILO082  
Ig kappa chain V region (2D3) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 16-Aug-1996  
C:Accession: PILO082  
R:Meek, K.; Haseman, C.; Pollok, B.; Alkan, S.S.; Bratt, M.; Slaoui, M.; Urbain, J.; Ca  
J. Exp. Med. 169, 519-533, 1989

```

A:Title: Structural characterization of antidiotypic antibodies; evidence that Ab2s
A:Reference number: PU0080; MUID:89094248
A:Accession: PU0082
A:Molecule type: mRNA
A:Residues: 1-106 <ME>
A:Experimental source: strain BALB/c
A:Note: the sequence shown here is from the V kappa region of an antidiotypic monoc
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match          38.0%; Score 482; DB 2; Length 106;
Best Local Similarity 87.4%; Pred. No. 6.1e-29;
Matches 90; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 5 ITOSPAILTSAPGKRWYTCRATPVSYSYMHYQOKPPSPKFWYTTSNLASGVPARFSG 64
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 LSGPAILTSAPGKRWYTCRATPVSYSYMHYQOKPPSPKFWYTTSNLASGVPARFSG 63

QY 65 GSGSTSYSLTVSRVYEAEDATYYCQWMSRSPDPFGGSKLEIK 107
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 64 GSGSTSYSLTVSRVYEAEDATYYCQWMSRSPDPFGGSKLEIK 106

RESULT 10
H32513
I9 heavy chain precursor V region (BXW16) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-May-1990 #sequence_rev:10 31-Dec-1990 #text_change 20-Mar-1998
C:Accession: H32513
R:Kotler, R.; Strichal, R.; Balderas, R.S.; Johnson, M.E.; Noonan, D.J.; Duchosal, M.A
J. Clin. Invest. 82, 852-860, 1988
A:Title: Immunoglobulin kappa light chain variable region gene complex organization a
A:Reference number: A94689; MUID:88331394
A:Accession: H32513
A:Molecule type: DNA
A:Residues: 1-137 <KOF>
A:Cross-references: GB:M20831; NID:g196949; PID:g196950
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match          37.9%; Score 480.5; DB 2; Length 137;
Best Local Similarity 74.4%; Pred. No. 1e-28;
Matches 93; Conservative 11; Mismatches 18; Indels 3; Gaps 1;

QY 113 SSKSESGKGVQLQDSGPVYVPGSGMKISCTSGYSTGHTMNVKSGKGNLEWIGLIN 172
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 15 AGVHE--IDLOOSGALVTRPGASVYKISCKASGYSFTGYNMNVKSGKSKLEWIGLIN 71

QY 173 PYNQDTNNOKFKGATFTVDKSSSTAYMELLSLSEDSAYYCARRTDWFEDWAGCT 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 PYGGTSTNNOKFKGATFTVDKSSSTAYMQLSLSEDSAYYCARKNYGSFDFWQGT 131

QY 233 TWTVS 237
   |||||
DB 132 TLTVS 136

RESULT 11
F45722
anti-glycoprotein H monoclonal antibody heavy-chain variable domain (Mab 33) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Sep-1993 #sequence_rev:10 18-Nov-1994 #text_change 09-May-1997
C:Accession: F45722
R:Simpson, J.A.; Chow, J.C.; Baker, J.; Avdalovic, N.; Yuan, S.; Au, D.; Co, M.S.; Va
J. Virol. 67, 489-496, 1993
A:Title: Neutralizing monoclonal antibodies that distinguish three antigenic sites on
A:Reference number: A45722; MUID:93100833
A:Accession: F45722
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-120 <SID>

```

A:Note: sequence extracted from NCBI backbone (NCBIF:120594)  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: glycoprotein

Query Match 37.6%; Score 477; DB 2; Length 120;  
Best Local Similarity 77.1%; Pred. No. 1.6e-28;  
Matches 91; Conservative 11; Mismatches 14; Indels 2; Gaps 2;

QY 122 VOLQSGPEVVPKPGSMKISCKTSGYFTGHTMNMVKOSHGNLEWIGL-IPYNGDTNY 180  
D 2 VOLQSGPELVKPGASVKISCKASGYFTGYTMNMVKOSHGNLEWIGLIPNNNGSY 61  
QY 181 NQKFGKATFTYDKSSSTAYMELLSLTSEDSAVYYCARR-VTDWTFDVGAGTIVTS 237  
D 62 NORFGKATLTLLDKSSSTAYMELLSLTSEDSAVYYCARRGFDYALDSWGQGSIVTS 119

## RESULT 12

Ig heavy chain V region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 19-Mar-1998 #sequence\_revision 19-Mar-1998 #text\_change 17-Apr-1998  
C:Accession: S26319  
R:Stark, S.E.; Caton, A.J.  
J:Exp. Med. 174, 613-624, 1991  
A:Title: Antibodies that are specific for a single amino acid interchange in a protein  
A:Reference number: S26309  
A:Accession: S26319  
A:Molecule type: mRNA  
A:Residues: 1-114 <STRA>  
A:Cross-references: EMBL:X59172  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin

Query Match 37.5%; Score 475; DB 2; Length 114;  
Best Local Similarity 78.4%; Pred. No. 2.1e-28;  
Matches 91; Conservative 9; Mismatches 10; Indels 6; Gaps 2;

QY 125 QESGPEVVPKPGSMKISCKTSGYFTGHTMNMVKOSHGNLEWIGLIPYNGDTNYNOKF 184  
D 1 QOSGPELVKPGASVKISCKASGYFTGYFMNMVKOSHGNLEWIGLRINPDRTYNOKF 60  
QY 185 KKGKATFTYDKSSSTAYMELLSLTSEDSAVYYCARRVTDWTF--DVGAGTIVTS 237  
D 61 KKGKATFTYDKSSSTAYMELLSLTSEDSAVYYCGK---DGYGAMDYWGQGSIVTS 113

## RESULT 13

PS0057  
Ig heavy chain precursor V region (PAR) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 08-Sep-1997  
C:Accession: PS0057  
R:Yasui, Y.; Takahashi, M.; Azuma, C.; Kanai, Y.; Honjo, T.  
J: Biochem. 104, 337-343, 1988  
A:Title: Biased expression of variable region gene families of the immunoglobulin heavy  
A:Reference number: PS0057; MUID:89197817  
A:Accession: PS0057  
A:Molecule type: DNA  
A:Residues: 1-135 <YAO>  
A:Cross-references: GB:D00307; NID:g220448; PID:d1000661; PID:g220449  
A:Note: the authors translated the codon AAG for residue 32 as Asn and GAC for 92 as Gly  
C:Comment: The gene encoding this protein was isolated from a hybridoma that produces an  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-135/Product: Ig heavy chain V region PAR #status predicted <MAT>

Query Match 37.5%; Score 474.5; DB 2; Length 135;  
Best Local Similarity 76.0%; Pred. No. 2.8e-28;

Matches 95; Conservative 9; Mismatches 16; Indels 5; Gaps 2;

QY 113 SKSSSEKGVOLQSGPEVVPKPGSMKISCKTSGYFTGHTMNMVKOSHGNLEWIGLIN 172  
D 15 AGVFESE--VOLQSGPELVKPGASVKISCKASGYFTGYFMNMVKOSHGNLEWIGLIN 71  
QY 173 PYNGDTNNQKFGKATFTYDKSSSTAYMELLSLTSEDSAVYYCARRVTDWTFDVGAGT 232  
D 72 PYNGDTFYNQKFGKATFTYDKSSSTAYMELLSLTSEDSAVYYCAR--LNYRGAYWGQGT 129

QY 233 TVTVS 237  
D 130 LVTVS 134

## RESULT 14

PH0099  
Ig heavy chain V region (anti-cyclosporin F) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 16-Aug-1996  
C:Accession: PH0099  
R:Schmitter, D.; Poch, O.; Zeder, G.; Heinrich, G.F.; Koehler, H.P.; Quesniaux, V.F.J.  
Mol. Immunol. 27, 1029-1038, 1990  
A:Title: Analysis of the structural diversity of monoclonal antibodies to cyclosporin  
A:Reference number: PH0087; MUID:91042649  
A:Accession: PH0099  
A:Molecule type: mRNA  
A:Residues: 1-119 <SCH>  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin  
F:31-35/Region: complementarity-determining 1  
F:50-66/Region: complementarity-determining 2  
F:99-106/Region: complementarity-determining 3

Query Match 37.4%; Score 473.5; DB 2; Length 119;  
Best Local Similarity 78.0%; Pred. No. 2.9e-28;  
Matches 92; Conservative 9; Mismatches 12; Indels 5; Gaps 2;

QY 122 VOLQSGPEVVPKPGSMKISCKTSGYFTGHTMNMVKOSHGNLEWIGLIPYNGDTNYN 181  
D 2 VOLQSGPELVKPGASVKISCKASGYFTGYTMNMVKOSHGNLEWIGLIPYNGDTNYN 61  
QY 182 QKFGKATFTYDKSSSTAYMELLSLTSEDSAVYYCARRVTDWTF--DVGAGTIVTS 237  
D 62 QKFGKATFTYDKSSSTAYMELLSLTSEDSAVYYCAR--GYTAAVDYWGQGSIVTS 116

## RESULT 15

F30502  
Ig heavy chain V region (A52) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Nov-1988 #sequence\_revision 03-Nov-1988 #text\_change 16-Aug-1996  
C:Accession: F30502  
R:Ellat, D.; Webster, D.M.; Rees, A.R.  
J: Immunol. 141, 1745-1753, 1988  
A:Title: V region sequences of anti-DNA and anti-RNA autoantibodies from NZB/NZW F-1  
A:Reference number: A30502; MUID:88315787  
A:Accession: F30502  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-119 <ELL>  
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology  
C:Keywords: heterotetramer; Immunoglobulin

Query Match 37.3%; Score 473; DB 2; Length 119;  
Best Local Similarity 77.1%; Pred. No. 3.1e-28;  
Matches 91; Conservative 11; Mismatches 14; Indels 2; Gaps 2;

QY 122 VOLQSGPEVVPKPGSMKISCKTSGYFTGHTMNMVKOSHGNLEWIGLIPYNGDTNYN 181  
D 2 VOLQSGPELVKPGASVKISCKASGYFTGYFMNMVKOSHGNLEWIGLRINPDRTYNOKF 61

Thu May 13 10:28:16 1999

us-08-704-178-3.rpr

Page 5

**Oy** 102 QKFGKATFYDKSSSTAYVELLSTSESAVYYCAR-RY-IDWTFDWGAGTIVTVS 237  
|||::|| |||||||: ||||||||| : || |||:||  
**D6** 62 QKFGQALTYDKSSSTAYMQLNSTSESAVYYCARGRLRGGYFDWGQTLLTVS 119

Search completed: May 13, 1999, 10:03:48  
Job time: 1380 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 09:55:33 ; Search time 16.53 Seconds

(without alignments)  
384.825 Million cell updates/sec

Title: US-08-704-178-3

Perfect score: 1267

Sequence: 1 MDLQITOSPALISAPGKRV.....RRVTDWTFDVGAGTTVTVS 237

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt\_36.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496.5	39.2	117	1	HV12_MOUSE
2	487.5	38.5	117	1	HV13_MOUSE
3	471	37.2	118	1	HV51_MOUSE
4	439.5	34.7	136	1	HV15_MOUSE
5	437	34.5	107	1	KV6E_MOUSE
6	432	34.1	107	1	KV6I_MOUSE
7	431	34.0	107	1	KV6G_MOUSE
8	431	34.0	107	1	KV6H_MOUSE
9	430	33.9	107	1	KV6J_MOUSE
10	429	33.9	108	1	KV6K_MOUSE
11	424.5	33.5	120	1	HV03_MOUSE
12	422	33.3	117	1	HV14_MOUSE
13	420	33.1	107	1	KV6D_MOUSE
14	419	33.1	107	1	KV6A_MOUSE
15	414	32.7	107	1	KV6B_MOUSE
16	412.5	32.6	139	1	HV07_MOUSE
17	412	32.5	107	1	KV6C_MOUSE
18	407	32.1	140	1	HV02_MOUSE
19	407	32.1	120	1	HV50_MOUSE
20	403.5	31.8	137	1	HV11_MOUSE
21	401	31.6	129	1	KV4A_MOUSE
22	397	31.3	107	1	KV6E_MOUSE
23	381.5	30.1	138	1	HV48_MOUSE
24	381	30.1	117	1	HV06_MOUSE
25	373.5	29.5	121	1	HV01_MOUSE
26	373	29.4	117	1	HV04_MOUSE
27	371	29.3	117	1	HV52_MOUSE
28	363.5	28.7	117	1	HV09_MOUSE
29	362	28.6	117	1	HV05_MOUSE
30	360.5	28.5	117	1	KV3T_MOUSE
31	360	28.4	111	1	HV49_MOUSE
32	355.5	28.1	111	1	KV3H_MOUSE
33	355.5	28.1	111	1	KV3R_MOUSE
34	355	28.0	117	1	HV10_MOUSE
35	353.5	27.9	111	1	KV3J_MOUSE
36	353.5	27.9	111	1	KV3L_MOUSE
37	352.5	27.8	131	1	KV3I_MOUSE
38	351.5	27.7	111	1	KV3O_MOUSE
39	347.5	27.4	108	1	KV1B_HUMAN
40	346.5	27.3	111	1	KV3M_MOUSE
41	346.5	27.3	111	1	KV3S_MOUSE
42	345.5	27.3	111	1	KV3K_MOUSE
43	345.5	27.3	111	1	KV3N_MOUSE

## ALIGNMENTS

44 344 27.2 109 1 KV3B\_HUMAN P01620 homo sapien  
45 343.5 27.1 111 1 KV3O\_MOUSE P01659 mus musculu

```
RESULT 1
ID HV12_MOUSE STANDARD; PRT; 117 AA.
AC P01756;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1991 (REL. 20, LAST ANNOTATION UPDATE)
DE IG HEAVY CHAIN V REGION (MOPC 104E).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN (1)
RP SEQUENCE.
RX MEDLINE; 83075344.
RA KERRY M.R., FUHRMAN J.S., SCHILLING J.W., ROGERS J., SIBLEY C.H.,
RA HOOD L.E.;
RL BIOCHEMISTRY 21:5415-5424(1982).
CC -1- THE SEQUENCE OF THE LIGHT CHAIN OF THIS IGM MYELOMA PROTEIN HAS
CC ALSO BEEN DETERMINED.
CC -1- THIS PROTEIN BINDS DEXTRAN.
DR PIR; A02039; MHMS4E.
DR HSSP; P01772; IFSV.
KW IMMUNOGLOBULIN V REGION; GLYCOPROTEIN.
FT DISULFID 22 BY SIMILARITY.
FT CARBOHYD 55
FT NON_TER 117
SQ SEQUENCE 117 AA; 12983 MW; 1F6CC304 CRC32;
```

Query Match 39.2%; Score 496.5; DB 1; Length 117;  
Best Local Similarity 81.0%; Pred. No. 5.6e-31;  
Matches 94; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

```
QY 122 VOLQSGPEVAVPFGSMKISCTSGTSGTGHMNVKQSHGNLEWIGLINPYNGDTN 181
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2 VOLQSGPELVKPGASVKSKASGTFDYMKVVKQSHGSLWIGLINPYNGDTN 61
QY 182 QKFKGATFTVKKSSSTAYMELSLTSEDSAYYCARVTDVFPVWGAGTTVTVS 237
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 62 QKFKGATFTVKKSSSTAYMELSLTSEDSAYYCARVTDVFPVWGAGTTVTVS 116

RESULT 2
ID HV13_MOUSE STANDARD; PRT; 117 AA.
AC P01757;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1991 (REL. 20, LAST ANNOTATION UPDATE)
DE IG HEAVY CHAIN V REGION (J558).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN (1)
RP SEQUENCE.
RX MEDLINE; 80078170.
RA SCHILLING J., CLEYINGER B., DAVIE J.M., HOOD L.;
RL NATURE 283:35-40(1980).
CC -1- THE SEQUENCES OF 10 HYBRIDOMA PROTEINS THAT ALSO BIND DEXTRAN
CC DIFFER FROM THAT SHOWN AT 1-7 POSITIONS, MANY OF WHICH OCCUR IN
CC THE D AND J SEGMENTS.
CC -1- THIS PROTEIN BINDS DEXTRAN.
DR PIR; A26242; MHMSJ5.
DR HSSP; P01789; ZFGW.
KW IMMUNOGLOBULIN V REGION.
FT DISULFID 22 BY SIMILARITY.
```

FT NON\_TER 117 117  
SQ SEQUENCE 117 AA: 13024 MM: E7548A05 CRC32;

Query Match  
Best Local Similarity 38.5%; Score 487.5; DB 1; Length 117;  
Matches 93; Conservative 9; Mismatches 13; Indels 1; Gaps 1;

OY 122 VOLQESGPEVVRKPGSMKISCKTSGYFTGHTNMVWKOSHGNLEWIGLINPYNGDTN 181  
DB 2 VOLQOGSPPELVKPGASVKSKASGYFTFDYKMWKQSHGKSLWIGIDINPNNGTSYN 61

OY 182 QKFKGATFTVDKSSSTAYMELLSTSEDSAVYICARRTDMYFDWGAAGTIVTS 237  
DB 62 QKFKGATFTVDKSSSTAYMQLNSTSEDSAVYICARD-RYMYFDWGAAGTIVTS 116

RESULT 3  
HV51\_MOUSE STANDARD; PRT; 118 AA.

AC P06330;  
DT 01-JAN-1988 (REL. 06, CREATED)  
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)  
DE IG HEAVY CHAIN V REGION (AC38 205.12).  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE: 84182519.  
RA DILPROP R., BOVENS J., SIEKEVITZ M., BEYREUTHER K., RAJESKY K.;  
RL EMO J. 3:517-523(1984).  
DR HSP: P01772; 1FGV.  
KW IMMUNOGLOBULIN V REGION.  
FT DOMAIN 1 98 V SEGMENT.  
FT DOMAIN 99 104 D SEGMENT.  
FT DOMAIN 105 118 J SEGMENT.  
FT DISULFID 22 96 BY SIMILARITY.  
FT NON\_TER 118 118  
SQ SEQUENCE 118 AA: 12934 MM: 2D1DC77 CRC32;

Query Match  
Best Local Similarity 37.2%; Score 471; DB 1; Length 118;  
Matches 91; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

OY 122 VOLQESGPEVVRKPGSMKISCKTSGYFTGHTNMVWKOSHGNLEWIGLINPYNGDTN 181  
DB 2 VOLQOGSPPELVKPGASVKISCKASGYFTFDYKMWKQSHGKSLWIGIDINPNNGTSYN 61

OY 182 QKFKGATFTVDKSSSTAYMELLSTSEDSAVYICARRTDMYFDWGAAGTIVTS 237  
DB 62 QKFKGATFTVDKSSSTAYMELRLSTSEDSAVYICARGYDFDWGAGTIVTS 117

RESULT 4  
HV15\_MOUSE STANDARD; PRT; 136 AA.

AC P01759;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)  
DE IG HEAVY CHAIN PRECURSOR V REGION (BC11).  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 82222262.  
RA KNAPP M.R., LIU C.-P., NEWELL N., WARD R.B., TUCKER P.W., STROBER S.,  
BEATYNER F.R.;

RL PROC. NATL. ACAD. SCI. U.S.A. 79:2996-3000(1982).  
DR EMBL: J00494; G195011; -.  
DR PIR: A02042; HWSB1.  
DR HSP: P01789; 1JEL.  
KW IMMUNOGLOBULIN V REGION; SIGNAL.  
FT SIGNAL 1 19  
FT CHAIN 20 136 IG HEAVY CHAIN V REGION (BC11).  
FT NON\_TER 136 136  
SQ SEQUENCE 136 AA: 15078 MM: E04F1C7F CRC32;

Query Match  
Best Local Similarity 34.7%; Score 439.5; DB 1; Length 136;  
Matches 84; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

OY 122 VOLQESGPEVVRKPGSMKISCKTSGYFTGHTNMVWKOSHGNLEWIGLINPYNGDTN 181  
DB 21 VOLQOGSPPEVVRPGSVKSKGSGYFTFDYKMWKQSHGKSLWIGISTYNGTSTYN 80

OY 182 QKFKGATFTVDKSSSTAYMELLSTSEDSAVYICARRTDMYFDWGAAGTIVTS 237  
DB 81 QKFKGATFTVDKSSSTAYMELARLSTSEDSANLYCARYGN-YFDYKMGCTIVTS 135

RESULT 5  
KV6F\_MOUSE STANDARD; PRT; 107 AA.

AC P04940;  
DT 13-AUG-1987 (REL. 05, CREATED)  
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE IG KAPPA CHAIN V-VI REGION (NO2-17.4.1).  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 83271467.  
RA KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;  
RL NATURE 304:320-324(1983).  
CC -I- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
DR EMBL: K00735; G196435; -.  
DR HSP: P01679; 1FVB.  
KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.  
FT DOMAIN 1 23 FRAMEWORK 1.  
FT DOMAIN 24 33 COMPLEMENTARITY-DETERMINING 1.  
FT DOMAIN 34 48 FRAMEWORK 2.  
FT DOMAIN 49 55 COMPLEMENTARITY-DETERMINING 2.  
FT DOMAIN 56 87 FRAMEWORK 3.  
FT DOMAIN 88 96 COMPLEMENTARITY-DETERMINING 3.  
FT DOMAIN 97 106 FRAMEWORK 4.  
FT DISULFID 23 87 BY SIMILARITY.  
FT NON\_TER 107 107  
SQ SEQUENCE 107 AA: 11561 MM: 1BC0C1E3 CRC32;

Query Match  
Best Local Similarity 34.5%; Score 437; DB 1; Length 107;  
Matches 80; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

OY 5 LTQSPALISASPEKVTMTCRATPSVYMHVYQKPGSPKPMIYTTSTLASGVPARFSG 64  
DB 4 LTQSPALISASGQVYTMCSASSSVYMHVYQKSGISPKMIYDTSLASGVPARFSG 63

OY 65 GSGSGISLTVSRVEADEAATYCCOWMSRSPPTFGGSKLEIK 107  
DB 64 SCSATSYSLITSMOAEADATYCCOWMSNPLTFAGTLEIK 106

RESULT 6  
KV6I\_MOUSE STANDARD; PRT; 107 AA.  
AC P04943;

```

DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE IG KAPPA CHAIN V-VI REGION (NO6-8.3.1).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83271467.
RA KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;
RL NATURE 304:320-324(1983).
CC -I- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.
DR EMBL: K00740; G196443; -.
DR HSSP: P01679; 1FVB.
KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.
FT DOMAIN 1 23 FRAMEWORK 1.
FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 3 48 FRAMEWORK 2.
FT DOMAIN 4 49 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 5 55 FRAMEWORK 3.
FT DOMAIN 6 56 FRAMEWORK 4.
FT DOMAIN 7 88 COMPLEMENTARITY-DETERMINING 3.
FT DOMAIN 8 97 FRAMEWORK 4.
FT DISULFID 97 106 BY SIMILARITY.
FT NON_TER 23 87
SQ SEQUENCE 107 AA; 11572 MW; 2D976197 CRC32;

*
Query Match
Best Local Similarity 76.7%; Score 432; DB 1; Length 107;
Matches 79; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 5 LTIOSPAILSASGEKVTMCRAIPSYVMHWYQOKGSPKPMIYTTSLASGVPARFSG 64
DB 4 LTIOSPAIMASPGOKVTMTCSASSSYVMHWYQOKSGTSPKRIYDTSLASGVPARFSG 63
QY 65 GSGSTSYSLTVSRVEADATYTCQOWSRSPPTFGGSKLEIK 107
DB 64 SSGATSYSLTISMQAEDATYTCQOWSNPLTFAGTLEIK 106

RESULT 7
KV6G_MOUSE STANDARD; PRT; 107 AA.
AC P04941;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE IG KAPPA CHAIN V-VI REGION (NO2-48.2.2).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83271467.
RA KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;
RL NATURE 304:320-324(1983).
CC -I- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.
DR EMBL: K00737; G196439; -.
DR HSSP: P01679; 1FVB.
KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.
FT DOMAIN 1 23 FRAMEWORK 1.
FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 3 44 FRAMEWORK 2.
FT DOMAIN 4 49 FRAMEWORK 3.
FT DOMAIN 5 56 FRAMEWORK 4.
FT DOMAIN 6 88 COMPLEMENTARITY-DETERMINING 3.
FT DOMAIN 7 97 FRAMEWORK 4.
FT DISULFID 97 106 BY SIMILARITY.
FT NON_TER 23 87
SQ SEQUENCE 107 AA; 11556 MW; 1AB8295D CRC32;

```

```

Query Match
Best Local Similarity 34.0%; Score 431; DB 1; Length 107;
Matches 79; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 5 LTIOSPAILSASGEKVTMCRAIPSYVMHWYQOKGSPKPMIYTTSLASGVPARFSG 64
DB 4 LTIOSPAIMASPGOKVTMTCSASSSYVMHWYQOKSGTSPKRIYDTSLASGVPARFSG 63
QY 65 GSGSTSYSLTVSRVEADATYTCQOWSRSPPTFGGSKLEIK 107
DB 64 SSGATSYSLTISMQAEDATYTCQOWSNPLTFAGTLEIK 106

RESULT 8
KV6H_MOUSE STANDARD; PRT; 107 AA.
AC P04942;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE IG KAPPA CHAIN V-VI REGION (NO5-61.1.2).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83271467.
RA KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;
RL NATURE 304:320-324(1983).
CC -I- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.
DR EMBL: K00739; G196443; -.
DR HSSP: P01679; 1FVB.
KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.
FT DOMAIN 1 23 FRAMEWORK 1.
FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 3 34 FRAMEWORK 2.
FT DOMAIN 4 49 FRAMEWORK 3.
FT DOMAIN 5 56 FRAMEWORK 4.
FT DOMAIN 6 88 COMPLEMENTARITY-DETERMINING 3.
FT DOMAIN 7 97 FRAMEWORK 4.
FT DISULFID 97 106 BY SIMILARITY.
FT NON_TER 23 87
SQ SEQUENCE 107 AA; 11605 MW; 1CF3CE8D CRC32;

Query Match
Best Local Similarity 34.0%; Score 431; DB 1; Length 107;
Matches 79; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 5 LTIOSPAILSASGEKVTMCRAIPSYVMHWYQOKGSPKPMIYTTSLASGVPARFSG 64
DB 4 LTIOSPAIMASPGOKVTMTCSASSSYVMHWYQOKSGTSPKRIYDTSLASGVPARFSG 63
QY 65 GSGSTSYSLTVSRVEADATYTCQOWSRSPPTFGGSKLEIK 107
DB 64 SSGATSYSLTISMQAEDATYTCQOWSNPLTFAGTLEIK 106

RESULT 9
KV6J_MOUSE STANDARD; PRT; 107 AA.
AC P04944;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE IG KAPPA CHAIN V-VI REGION (NO5-78.2.6).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83271467.
RA KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;

```

RL NATURE 304:320-324(1983).  
 CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
 DR EMBL: K00744; G196453; -.  
 DR HSSP: P01679; 1FVB.  
 KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.  
 FT DOMAIN 1 23 FRAMEWORK 1.  
 FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.  
 FT DOMAIN 3 44 FRAMEWORK 2.  
 FT DOMAIN 4 49 55 COMPLEMENTARITY-DETERMINING 2.  
 FT DOMAIN 5 56 87 FRAMEWORK 3.  
 FT DOMAIN 6 88 96 COMPLEMENTARITY-DETERMINING 3.  
 FT DOMAIN 7 97 106 FRAMEWORK 4.  
 FT DISULFID 23 87 BY SIMILARITY.  
 FT NON\_TER 107  
 SQ SEQUENCE 107 AA; 11611 MW; E926BBEB CRC32;

Query Match 33.9%; Score 430; DB 1; Length 107;  
 Best Local Similarity 76.7%; Pred. No. 4.7e-26;  
 Matches 79; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 5 LTGSPAILASPEKVTMTCRATPSYSYMHVYQKPGSSPKPIYTTSNLASGVPARFSG 64  
 DB 4 LTGSPAILASPEKVTMTCSASSYSYMHVYQKSGTSPKRIYDTSKLASGVPARFSG 63  
 QY 65 GSGSTSYSLTVSRVEADATYYCOQWSSRPPTFGGSKLEIK 107  
 DB 64 GSGSTSYSLTITSMQADATYYCOQWSSNPLTFGSGTKLEIK 106

RESULT 10  
 ID KV6K\_MOUSE STANDARD; PRT; 108 AA.  
 AC P04945;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
 DE IG KAPPA CHAIN V-VI REGION (NO2-6.1).  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 83271467.  
 RA KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;  
 RL NATURE 304:320-324(1983).  
 DR EMBL: K00746; G554059; -.  
 DR HSSP: P01679; 1BAE.  
 KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.  
 FT DOMAIN 1 23 FRAMEWORK 1.  
 FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.  
 FT DOMAIN 3 44 FRAMEWORK 2.  
 FT DOMAIN 4 49 55 COMPLEMENTARITY-DETERMINING 2.  
 FT DOMAIN 5 56 87 FRAMEWORK 3.  
 FT DOMAIN 6 88 98 COMPLEMENTARITY-DETERMINING 3.  
 FT DOMAIN 7 108 108 FRAMEWORK 4.  
 FT DISULFID 23 87 BY SIMILARITY.  
 FT NON\_TER 108  
 SQ SEQUENCE 108 AA; 11713 MW; 8B48A204 CRC32;

Query Match 33.9%; Score 429; DB 1; Length 108;  
 Best Local Similarity 77.1%; Pred. No. 5.7e-26;  
 Matches 81; Conservative 11; Mismatches 11; Indels 2; Gaps 1;

QY 5 LTGSPAILASPEKVTMTCRATPSYSYMHVYQKPGSSPKPIYTTSNLASGVPARFSG 64  
 DB 4 LTGSPAILASPEKVTMTCSASSYSYMHVYQKPGSSPKRIYDTSKLASGVPARFSG 63  
 QY 65 GSGSTSYSLTVSRVEADATYYCOQWSSRP--TFGGGSKLEIK 107  
 DB 64 GSGSTSYSLTITSMQADATYYCOQWSSYPPMLTFGAGTKLEIK 108

RESULT 11  
 ID HV03\_MOUSE STANDARD; PRT; 120 AA.  
 AC P01747;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 21-JUL-1986 (REL. 01, LAST ANNOTATION UPDATE)  
 DE IG HEAVY CHAIN V REGION (36-65).  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 83131846.  
 RA SIEKEVITZ M., GEFTER M.L., BRODEUR P., RIBLET R.,  
 RA MARSHAK-ROTHSTEIN A.;  
 RL EUR. J. IMMUNOL. 12:1023-1032(1982).  
 CC -1- FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER DIFFERENTIATED GENES  
 CC THAT HYBRIDIZE TO THIS ONE, THE AUTHORS CONCLUDE THAT ALL OF  
 CC THESE V REGIONS HAVE REARRANGED TO THE SAME J SEGMENT, JH2.  
 DR PIR: A02028; HVM8G7.  
 DR HSSP: P01789; 6FAB.  
 KW IMMUNOGLOBULIN V REGION; ANTIBODY; HYBRIDOMA.  
 FT NON\_TER 120  
 SQ SEQUENCE 120 AA; 13307 MW; BBA8CCAI CRC32;

Query Match 33.5%; Score 424.5; DB 1; Length 120;  
 Best Local Similarity 68.1%; Pred. No. 1.4e-25;  
 Matches 81; Conservative 14; Mismatches 21; Indels 3; Gaps 1;

QY 122 VOLOSPEVTVRGSGMKISCKTSGSYFGHTMNVKSHGKLEIGLINFNGDTN 181  
 DB 1 VOLOSAGELVVRAGSSVYKASCSAGTFTSYGINVWKPQPGCLERIGINGNGYTKYN 60  
 QY 182 QRFKGRATFTVDSSTAYMELLSLTSBDAVYCARV--TDWYFDVWAGATVTVS 237  
 DB 61 EKFKGRTLVTKSSSTAVMQRSLTSBDAVYFCARSYVGGSYVFDWGGTTLTVS 119

RESULT 12  
 ID HV14\_MOUSE STANDARD; PRT; 117 AA.  
 AC P01758;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE IG HEAVY CHAIN PRECURSOR V REGION (108A).  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 81245215.  
 RA GIVOL D., ZAKUT R., EFFRON K., RECHAVI G., RAM D., COHEN J.B.;  
 RL NATURE 292:426-430(1981).  
 DR EMBL: J00488; G554028; -.  
 DR PIR: A02041; HVM8A.  
 DR HSSP: P01789; 11G1.  
 KW IMMUNOGLOBULIN V REGION; SIGNAL.  
 FT SIGNAL 1 19  
 FT CHAIN 20 117 IG HEAVY CHAIN V REGION (10A).  
 FT NON\_TER 117  
 SQ SEQUENCE 117 AA; 12972 MW; 941C002E CRC32;

Query Match 33.3%; Score 422; DB 1; Length 117;  
 Best Local Similarity 75.5%; Pred. No. 2.1e-25;  
 Matches 80; Conservative 8; Mismatches 16; Indels 0; Gaps 0;  
 QY 113 SGKSGRGVLOESGPEVTVRGSGMKISCKTSGSYFGHTMNVKSHGKLEIGL 172  
 DB 113 SGKSGRGVLOESGPEVTVRGSGMKISCKTSGSYFGHTMNVKSHGKLEIGL 172





















GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 09:37:42 ; Search time 23.53 Seconds

(without alignments)  
203.715 Million cell updates/sec

Title: US-08-704-178-3

Perfect score: 1267

Sequence: 1 MDQLTQSPAILLSASPGKRV.....RRYTDWTFDVGAGTTVTS 237

Scoring table: BL0SUM62

Searched: 162890 seqs, 20225328 residues

Database : A\_Geneseq\_34:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1267	100.0	237	1	R94020	Anti-erbB2 scFv. I
2	1267	100.0	237	1	W15185	Single-chain anti-
3	1256	99.1	236	1	R45442	Sequence of the si
4	883.5	69.7	240	1	W25784	Anti-gp130 antibod
5	863	68.1	244	1	R60780	Fv(GP-2) immunosup
6	842	66.5	302	1	R60206	Bispecific CD3-16f
7	795.5	62.8	239	1	W15186	Single-chain anti-
8	793.5	62.6	238	1	R45443	Sequence of the si
9	784.5	61.9	243	1	R60781	Fv(GP-4) immunosup
10	782.5	61.8	301	1	W37085	Anti-human SC sing
11	748	59.0	272	1	R34672	26-10 scFv. Express
12	746	58.9	251	1	P80152	Multifunctional pr
13	745	58.8	239	1	W02191	18-2-3/TRY202' sin
14	742	58.6	239	1	R43679	Single chain polyp
15	742	58.6	239	1	R99649	Single chain bindi
16	738	58.2	241	1	R06482	18-2-3-/TRY202'. S
17	736.5	58.1	486	1	R37649	Sequence of divale
18	736.5	58.1	262	1	R97380	CC49/212 SCA PLAP
19	736.5	58.1	264	1	R97381	PLAP CC49/212 SCA
20	734	57.9	241	1	R37645	Sequence of a mult
21	731.5	57.7	250	1	R43678	Single chain polyp
22	731.5	57.7	250	1	R99648	Single chain bindi
23	731.5	57.7	250	1	W02190	TR104b single cha
24	730	57.6	271	1	R20185	Sequence of the sh
25	726.5	57.3	246	1	R43677	Single chain polyp
26	722.5	57.0	246	1	R99647	Single chain bindi
27	722	57.0	473	1	W26646	Chimeric receptor
28	722	57.0	473	1	W26647	Chimeric receptor
29	722	57.0	403	1	W26648	Chimeric receptor
30	722	57.0	651	1	W26649	Chimeric receptor
31	722	57.0	692	1	W26650	Chimeric receptor
32	722	57.0	382	1	W26651	Chimeric receptor
33	715	56.4	238	1	R55864	4-4-20 VL / 217 /
34	713.5	56.3	250	1	R06481	TR104b. Single po
35	710	56.0	238	1	R37650	Sequence of single
36	705.5	55.7	248	1	R77616	Humanised CDR-graf
37	704.5	55.6	242	1	R43680	Single chain polyp
38	704.5	55.6	242	1	R99650	Single chain bindi
39	704.5	55.6	242	1	W02192	18-2-3/TRY59 singl
40	699.5	55.2	248	1	R77605	Murine 5G1.1M1 scF
41	698.5	55.1	246	1	W02189	TRX59 single chain
42	694.5	54.8	242	1	R06483	18-2-3-/TRY59. Sin
43	693	54.7	272	1	W00557	Nematode salivary

ALIGNMENTS

44 692.5 54.7 248 1 R77607 Humanised CDR-graf  
45 689 54.4 272 1 W43913 Mus musculus antib

RESULT	ID	Query Match	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	R94020	100.0%	1267	1	237	237	237	0	0	0
2	R94020	100.0%	1267	1	237	237	237	0	0	0
3	R94020	100.0%	1267	1	237	237	237	0	0	0
4	R94020	100.0%	1267	1	237	237	237	0	0	0
5	R94020	100.0%	1267	1	237	237	237	0	0	0
6	R94020	100.0%	1267	1	237	237	237	0	0	0
7	R94020	100.0%	1267	1	237	237	237	0	0	0
8	R94020	100.0%	1267	1	237	237	237	0	0	0
9	R94020	100.0%	1267	1	237	237	237	0	0	0
10	R94020	100.0%	1267	1	237	237	237	0	0	0
11	R94020	100.0%	1267	1	237	237	237	0	0	0
12	R94020	100.0%	1267	1	237	237	237	0	0	0
13	R94020	100.0%	1267	1	237	237	237	0	0	0
14	R94020	100.0%	1267	1	237	237	237	0	0	0
15	R94020	100.0%	1267	1	237	237	237	0	0	0
16	R94020	100.0%	1267	1	237	237	237	0	0	0
17	R94020	100.0%	1267	1	237	237	237	0	0	0
18	R94020	100.0%	1267	1	237	237	237	0	0	0
19	R94020	100.0%	1267	1	237	237	237	0	0	0
20	R94020	100.0%	1267	1	237	237	237	0	0	0
21	R94020	100.0%	1267	1	237	237	237	0	0	0
22	R94020	100.0%	1267	1	237	237	237	0	0	0
23	R94020	100.0%	1267	1	237	237	237	0	0	0
24	R94020	100.0%	1267	1	237	237	237	0	0	0
25	R94020	100.0%	1267	1	237	237	237	0	0	0
26	R94020	100.0%	1267	1	237	237	237	0	0	0
27	R94020	100.0%	1267	1	237	237	237	0	0	0
28	R94020	100.0%	1267	1	237	237	237	0	0	0
29	R94020	100.0%	1267	1	237	237	237	0	0	0
30	R94020	100.0%	1267	1	237	237	237	0	0	0
31	R94020	100.0%	1267	1	237	237	237	0	0	0
32	R94020	100.0%	1267	1	237	237	237	0	0	0
33	R94020	100.0%	1267	1	237	237	237	0	0	0
34	R94020	100.0%	1267	1	237	237	237	0	0	0
35	R94020	100.0%	1267	1	237	237	237	0	0	0
36	R94020	100.0%	1267	1	237	237	237	0	0	0
37	R94020	100.0%	1267	1	237	237	237	0	0	0
38	R94020	100.0%	1267	1	237	237	237	0	0	0
39	R94020	100.0%	1267	1	237	237	237	0	0	0
40	R94020	100.0%	1267	1	237	237	237	0	0	0
41	R94020	100.0%	1267	1	237	237	237	0	0	0
42	R94020	100.0%	1267	1	237	237	237	0	0	0
43	R94020	100.0%	1267	1	237	237	237	0	0	0

KM Immunodiagnosis; treatment; cytotoxic agent; erbB-2.  
 OS Mus musculus.  
 OS Synthetic.  
 FH Key  
 FT region 108..121  
 FT /label= linker  
 FT US5587458-A.  
 FN 24-DEC-1996.  
 PD 07-OCT-1991; 772270.  
 PR 07-OCT-1991; US-772270.  
 PR 30-JUN-1992; US-906555.  
 PR 14-MAY-1993; US-061092.  
 PA (ARON-) ARONEX PHARM INC.  
 PI Bird RE, Kasprzyk PG, King CR;  
 DR WPI; 97-064831/06.  
 DR N-PSDB; T65006.  
 PT Single chain antibodies specific for erbB-2 protein, gp185 - with  
 PT labels or cytotoxin, useful for detection and treatment of tumour  
 PT cells expressing this protein  
 PS Claim 2; Columns 25-28; 28pp; English.  
 CC The present sequence represents a claimed single-chain antibody, e23  
 CC designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23  
 CC was generated by immunising mice with N/erbB-2 cells overexpressing  
 CC the gp185 protein, removing spleen cells and producing hybridomas  
 CC by standard techniques. Messenger RNA coding for the anti-erbB-2  
 CC monoclonal antibody was isolated and converted to cDNA. Regions  
 CC coding for the heavy- and light- chain variable regions were then  
 CC amplified by PCR and joined via a sequence encoding a peptide  
 CC linker. The resulting single-chain antibody is useful for in vitro  
 CC diagnosis of tumour cells which overexpress the erbB-2 gp185  
 CC marker, e.g. breast, ovarian and non-small cell lung carcinomas,  
 CC and, when coupled to a cytotoxic agent, to treat such tumours.  
 SQ Sequence 237 AA;

Query Match 100.0%; Score 1267; DB 1; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-80;  
 Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLQTSPPAILASPGKVTMTCRATPSVSYMHVYQKPGSSPKWITTTNLSAGVPA 60  
 DB 1 MDLQTSPPAILASPGKVTMTCRATPSVSYMHVYQKPGSSPKWITTTNLSAGVPA 60  
 QY 61 FSGGGSGTSTLYSRVDAEDATYYCCQMSRSPPTFGGSKLEIKGSTSGSKSSECK 120  
 DB 61 FSGGGSGTSTLYSRVDAEDATYYCCQMSRSPPTFGGSKLEIKGSTSGSKSSECK 120  
 QY 121 GVQLOESGEEVYKPGSKMISCKTSGTFTGHTMMVVKSHKKNLEWIGLIMPYNGDTMY 180  
 DB 121 GVQLOESGEEVYKPGSKMISCKTSGTFTGHTMMVVKSHKKNLEWIGLIMPYNGDTMY 180  
 QY 181 NQFKKATFTYDKSSSTAYMELSLTSDSAVYYCARRVMTWYFDWAGATTYVS 237  
 DB 181 NQFKKATFTYDKSSSTAYMELSLTSDSAVYYCARRVMTWYFDWAGATTYVS 237

RESULT 3  
 R45442  
 ID R45442 standard; Protein; 236 AA.  
 AC R45442.  
 DT 21-JUL-1994 (first entry)  
 DE Sequence of the single chain anti-erbB2 antibody, Ab no.23.  
 KW Single chain anti-erbB1 antibody; cancer therapy; prevention;  
 KW monoclonal antibody.  
 OS Synthetic.  
 FN WO9400136-A.  
 PD 06-JAN-1994.  
 PR 21-OCT-1992; 008545.  
 PR 30-JUN-1992; US-906555.  
 PA (MOLE-) MOLECULAR ONCOLOGY INC.  
 PI Kasprzyk PG, King CR;  
 DR WPI; 94-025878/03.  
 DR N-PSDB; Q55180.

PT Treatment of malignancies over-expressing ERB-12 - using at least  
 PT 2 monoclonal antibodies which recognise different epitopes on  
 PT gp185  
 PS Example; Fig 7; 37pp; English.  
 CC The source of human erbB-2 protein for the prodn. of antibodies no.  
 CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell  
 CC engineered to express the human erbB-2 protein on its surface (N/  
 CC erbB-2). Abs no. 21 and 23 are directed against the extracellular  
 CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly  
 CC growing tumours were used in a trial of the efficacy of the Abs. In  
 CC animals given a combination of the 2 Abs, tumours completely  
 CC regressed after 11 days.  
 SQ Sequence 236 AA;

Query Match 99.1%; Score 1256; DB 1; Length 236;  
 Best Local Similarity 99.2%; Pred. No. 1.9e-79;  
 Matches 234; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 DLQLOSPAILASPGKVTMTCRATPSVSYMHVYQKPGSSPKWITTTNLSAGVPA 61  
 DB 1 DLQLOSPAILASPGKVTMTCRATPSVSYMHVYQKPGSSPKWITTTNLSAGVPA 60  
 QY 62 FSGGGSGTSTLYSRVDAEDATYYCCQMSRSPPTFGGSKLEIKGSTSGSKSSECK 121  
 DB 61 FSGGGSGTSTLYSRVDAEDATYYCCQMSRSPPTFGGSKLEIKGSTSGSKSSECK 120  
 QY 122 VOLOESGEEVYKPGSKMISCKTSGTFTGHTMMVVKSHKKNLEWIGLIMPYNGDTMYN 181  
 DB 121 VOLOESGEEVYKPGSKMISCKTSGTFTGHTMMVVKSHKKNLEWIGLIMPYNGDTMYN 180  
 QY 182 QFKKATFTYDKSSSTAYMELSLTSDSAVYYCARRVMTWYFDWAGATTYVS 237  
 DB 181 QFKKATFTYDKSSSTAYMELSLTSDSAVYYCARRVMTWYFDWAGATTYVS 236

RESULT 4  
 W25784  
 ID W25784 standard; Protein; 240 AA.  
 AC W25784.  
 DT 22-DEC-1997 (first entry)  
 DE Anti-gp130 antibody derived scFv.  
 KW Single chain Fv; scFv; antibody; E. coli; anti-T3 antibody;  
 KW anti-gp130 antibody; inclusion body; chaperonin.  
 OS Synthetic.  
 FN J09220092-A.  
 PD 26-AUG-1997.  
 PR 15-FEB-1996; 027622.  
 PR 15-FEB-1996; JP-027622.  
 PA (TOYU) TOSOH CORP.  
 DR WPI; 97-474306/44.  
 DR N-PSDB; T91615.  
 PT Producing single chain Fv antibody in Escherichia coli - by  
 PT expression in an inclusion body, followed by protein folding or by  
 PT co-expression with a chaperonin as a soluble fraction  
 PS Claim 4; Page 6-7; 9pp; Japanese.  
 CC The sequences given in W25783-84 represent single chain Fv (scFv)  
 CC antibodies which are produced in E.coli. The scFv's are derived  
 CC from an anti-T3 antibody or an anti-gp130 antibody and are produced  
 CC either by: (1) expression as an inclusion body, followed by folding  
 CC (i.e. by denaturation and solubilisation) or (2) expression as a  
 CC soluble fraction by co-expression with chaperonin. Using the methods,  
 CC scFv can be produced in E. coli, either in a soluble fraction or in  
 CC inclusion bodies. Like the parent double chain antibodies, the scFv  
 CC polypeptides are useful in immunological diagnosis or for separation  
 CC and purification, but they lessen the side effects caused by antibody  
 CC constant regions. The availability of scFv polypeptides contributes  
 CC to fundamental research and development of diagnostic and therapeutic  
 CC drugs.  
 SQ Sequence 240 AA;

Query Match 69.7%; Score 883.5; DB 1; Length 240;



DB 263 GAGTCTCVS 271

RESULT 7  
ID W15186  
AC W15186 standard; Protein; 239 AA.  
DE Single-chain anti-erbB2 antibody e21(Fv).  
KW Single chain antibody; variable region; light chain; heavy chain;  
KW breast cancer; ovarian cancer; non-small cell lung carcinoma;  
KW immunodiagnosis; treatment; cytotoxic agent; erbB-2.  
OS Mus musculus.  
OS Synthetic.  
FH Key  
FT region  
FT 106.119  
FT /label-linker  
PN US5587458-A.  
PD 24-DEC-1996.  
PF 07-OCT-1991; 772270.  
PR 07-OCT-1991; US-772270.  
PR 30-JUN-1992; US-906555.  
PR 14-MAY-1993; US-061092.  
PA (ARON-) ARONEX PHARM INC.  
PI Bird RE, Kasprzyk PG, King CR;  
PI WPI: 97-064831/06.  
DR N-PSDB: T65007.  
PT Single chain antibodies specific for erbB-2 protein, gp185 - with  
PT labels or cytotoxin, useful for detection and treatment of tumour  
PT cells expressing this protein.  
PS Claim 2: Columns 27-30; 28pp; English.  
CC The present sequence represents a claimed single-chain antibody,  
CC designated e21(Fv), which binds to erbB-2. Monoclonal antibody e21  
CC was generated by immunising mice with N/erbB-2 cells overexpressing  
CC the gp185 protein, removing spleen cells and producing hybridomas  
CC by standard techniques. Messenger RNA coding for the anti-erbB-2  
CC monoclonal antibody was isolated and converted to cDNA. Regions  
CC coding for the heavy- and light- chain variable regions were then  
CC amplified by PCR and joined via a sequence encoding a peptide  
CC linker. The resulting single-chain antibody is useful for in vitro  
CC diagnosis of tumour cells which overexpress the erbB-2 gp185  
CC marker, e.g. breast, ovarian and non-small cell lung carcinomas,  
CC and, when coupled to a cytotoxic agent, to treat such tumours.  
SQ Sequence 239 AA;

Query Match 62.8%; Score 795.5; DB 1; Length 239;

Best Local Similarity 64.3%; Pred. No. 6.9e-48;  
Matches 155; Conservative 32; Mismatches 45; Indels 9; Gaps 2;

QY 3 LQLOSPALISASGKGVYTCRATPSVSYMHYQOKPGSSPKPMIYTTNLSAGVPARF 62  
DB 1 MQLTQSPALISASGKGVYTCRATPSVSYMHYQOKPGSSPKPMIYTTNLSAGVPARF 60  
QY 63 SGGSGSYSLTVSRVAEDAAATYYCOQWSRSPPTFGSGKLEIKSGTSSGKSSSEKGV 122  
DB 61 SGGSGSYSLTVSRVAEDAAATYYCOQWSRSPPTFGSGKLEIKSGTSSGKSSSEKGV 120  
QY 123 QLOESGEVYKPGGSMKISCKTSGYSTGHTMMVYKQSHKKNLEWIGLINPYNGDTNYNQ 182  
DB 121 QLOOSGVELVKGALVYKLSKASDENIKDYIHVYKQRPQGLEWIGIMHPENGNTVYDP 180  
QY 183 KFKGKAFTYDKSSSTAYMELLSITSEDSAVYVCARVTDWYFDV-----WGAGTIVY 236  
DB 181 KFGKASITADTSSNAAYLQLSITSEDTAVYCA---SYIYSAYYAMYWGOGISVTV 237  
QY 237 S 237  
DB 238 S 238

RESULT 8  
R45443.

ID R45443 standard; Protein; 238 AA.

AC R45443;  
DE 21-JUL-1994 (first entry)  
DE Sequence of the single chain anti-erbB2 antibody, Ab no.21.  
KW Single chain anti-erbB1 antibody; cancer therapy; prevention;  
KW monoclonal antibody.  
OS Synthetic.  
PN WO9400136-A.  
PD 06-JAN-1994.  
PF 21-OCT-1992; 008545.  
PR 30-JUN-1992; US-906555.  
PA (MOLE-) MOLECULAR ONCOLOGY INC.  
PI Kasprzyk PG, King CR;  
PI WPI: 94-025878/03.  
DR N-PSDB: Q55181.  
PT Treatment of malignancies over-expressing ERB-12 - using at least  
PT 2 monoclonal antibodies which recognise different epitopes on  
PT gp185  
PS Example: Fig 8; 37pp; English.  
CC The source of human erbB-2 protein for the prodn. of antibodies no.  
CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell  
CC engineered to express the human erbB-2 protein on its surface (N/  
CC erbB-2). Abs no. 21 and 23 are directed against the extracellular  
CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly  
CC growing tumours were used in a trial of the efficacy of the Abs. In  
CC animals given a combination of the 2 Abs, tumours completely  
CC regressed after 11 days.  
SQ Sequence 238 AA;

Query Match 62.6%; Score 793.5; DB 1; Length 238;

Best Local Similarity 64.6%; Pred. No. 9.5e-48;  
Matches 155; Conservative 31; Mismatches 45; Indels 9; Gaps 2;

QY 4 QLTQSPALISASGKGVYTCRATPSVSYMHYQOKPGSSPKPMIYTTNLSAGVPARF 63  
DB 1 QLTQSPALISASGKGVYTCRATPSVSYMHYQOKPGSSPKPMIYTTNLSAGVPARF 60  
QY 64 GGGSGSYSLTVSRVAEDAAATYYCOQWSRSPPTFGSGKLEIKSGTSSGKSSSEKGV 123  
DB 61 GGGSGSYSLTVSRVAEDAAATYYCOQWSRSPPTFGSGKLEIKSGTSSGKSSSEKGV 120  
QY 124 LOESGEVYKPGGSMKISCKTSGYSTGHTMMVYKQSHKKNLEWIGLINPYNGDTNYNQ 183  
DB 121 LOOSGVELVKGALVYKLSKASDENIKDYIHVYKQRPQGLEWIGIMHPENGNTVYDPK 180  
QY 184 KFKGKAFTYDKSSSTAYMELLSITSEDSAVYVCARVTDWYFDV-----WGAGTIVY 237  
DB 181 FQKASITADTSSNAAYLQLSITSEDTAVYCA---SYIYSAYYAMYWGOGISVTVS 237  
RESULT 9  
ID R60781 standard; Protein; 243 AA.  
AC R60781;  
DE 09-MAY-1995 (first entry)  
DE Fv(GP-4) immunosuppressive.  
KW Mab: monoclonal antibody; hybridoma; interleukin-2; IL-2;  
KW Fv; antibody variable region; GP-4; Fv(GP-4);  
KW immunosuppressive.  
OS Mus sp.  
PN EP-621338-A.  
PD 26-OCT-1994.  
PF 21-APR-1994; 106257.  
PR 21-APR-1993; JP-094491.  
PR 07-MAR-1994; JP-036065.  
PA (AJIN ) AJINOMOTO KK.  
PI Hamura J, Kanayama Y, Nakazawa H, Shimamura T, Sugamura K;  
PI Takeshita T;  
PI WPI: 94-325948/41.  
DR N-PSDB: Q73679.  
PT Immunosuppressant polypeptide - has ability to block  
PT interleukin-2 response

PS Claim 9: Page 27-28; 37p: English.  
CC MAB capable of binding to the gamma chain of the IL-2 receptor, and  
CC thus of blocking the IL-2 response, is produced by mouse hybridoma  
CC line GP-4 (EERM BP-4660). D. encoding the variable region of  
CC this MAB was expressed in *E. coli*, yielding Fv(GP-4) with  
CC immunosuppressive activity.  
CC Sequence 243 AA:  
50

Query Match	61.94;	Score 784.5;	DB 1;	length 243;
Best Local Similarity	63.54;	Pred. No. 4e-47;		
Matches 155;	Conservative 32;	Mismatches 48;	Indels 9;	Gaps 4;
QY	1 MDLQITSPALISASPGKRVWTTCRATPSY-SYMHWYQOKPPSSPKRWIYTTSTNLASGVP	59		
	:                       : : :           :			
Db	1 MDLILTQSPALISASVGEIVTITCGASGNIHNYLAMYQOKQKRSQLLVYNNKTLADGVP	60		
QY	60 ARFSGGGSGTISYSLTVSRVEADAAATYYCCQONSRRSPPTFFGGSKLEIK-GSTSGSGKSE	118		
	:       :       : : :       : :             :       :			
Db	61 SRFSGSGSGTQYSLKINSIQLPEDFSGYCQHFWSPPTWTFGGGTKEIDLKVKKSSGSGSEK	120		
QY	119 GKQVQLDSSGPEPVAVPGSGMKTISCTSGYTFGHIMNWKSKHGNLEWIGLINIPNYCOT	178		
	:   :			
Db	121 STQVLEKSGPELVKQASVYKISCSKASGSEFPGYVHWVWKSHVKSLEWIGRIINPGAT	180		
QY	179 NYNQKFKKATFTVVKSSSTAYMELSLTSEDSAYYYCAKRTIDW-----EDVWAGAT	233		
	:     :     :                               : :			
Db	181 SYNQFKKAKSLTVDKSSSTAYMELSLTSEDSAYYYCARE--NYGSSYGFAYWQGGAL	238		
QY	234 VTVS 237			
Db	239 VTVS 242			

CC is delivered to a cell surface via the SECY binding moiety. The nucleic acid can be compacted at high concentrations with the carrier molecule at a critical salt concentration. The condensation of such complexes provides structural features to the DNA/cationic lipid complex that CC prolongs in vivo expression.

[illegible]

CC comprising only the VH or VL regions, and single chain molecules may  
CC be used to specifically bind one or more of the same antigens as the  
CC full length antibody from which they are derived.  
SQ Sequence 272 AA;

**SQ Sequence 272 AA;**

Query Match	59.0%	Score 748	DB 1	Length 272
Best Local Similarity	60.28%	Pred. No.	1.4e-44	
Matches 148	Conservative 33	Mismatches 55	Indels 10	Gaps 4

QY	2	DLQLOTPSAIIASAPGEVMTWCRAATPSV-----SYMHVQOQPGSGSPKPMWITTTNLA	55
		: : :     : : :   : :   :   :   :   :   :   :	
Db	26	DVMTQTFPLSLPYELGLGDASISCRSSQSLVHNSONTYLTNWLQKAAGSPKLLITKYSNRF	85
QY	56	SGVPARESGGSGSTSYSLTYSRVAEADANATYYCOQWERSPPTFGGSKLEIK--GSTS-G	112
		: : :   : :   : :   : :   : :   : :   : :   : :	
Db	86	SGVDREFGSGSGSDFTLKTISRVAEEDLGITFCQOTHPVPTFGGTRKLEIKRGGSGSGG	145
QY	113	SGKSEGGVQLOESGPEVNVPGGSGMKISCTSGTSGTGHMMNVKROSHGKNLEWIGLIN	172
		: :   : :   : :   : :   : :   : :   : :   : :	
Db	146	SGSGSGSGSEVQLQOOSGPELVNPGASVNRKSSCSGITTTDFPMNVVRSHGKSLDIYGIS	205
QY	173	PLYNGDTINQKFKKATFTVDKSSSTAYMELLSTLSDSAAYYCA--RRVTDWTFYDWAG	231
		: :   : :   : :   : :   : :   : :   : :   : :	
Db	206	PYSGVTGYNQKFKKATFTLVKSSSTAYMELRSLTSEDSAYYCAAGSGGNMAMDYWGHC	265
QY	232	TTTVTS 237	
		: :   : :	
Db	266	ASVTVS 271	

**RESULT 12**

ID P80152 standard; protein; 251 AA.

DT 20-OCT-1990 (first entry)  
DE Multifunctional protein.  
KW Multifunctional protein; biosynthetic antibody binding site;  
KM murine monoclonal 26-10.  
PN MO809344-A.  
PD 01-DEC-1988.  
PE 19-MAY-1988; U01737.  
PR 21-MAY-1987; US-052800.  
PA (CREA) Creative Biomolecules Inc.  
PI Huston US, Oppermann H;  
DR WPI; 88-353928/49.  
DR N-PSDB; N80178.  
PT Recombinant multifunctional protein - having antibody binding site and a  
PT sequence for biological activity, ion sequestering or binding to a solid  
PT support.  
PS Disclosure; 15pp; English.  
CC The multifunctional protein comprises a single chain biosynthetic Ab  
CC binding site (comprising sequences mimicking the light chain and heavy  
CC chain variable regions of murine monoclonal 26-10) with the specificity  
CC of murine monoclonal 26-10, linked through a spacer to the FB fragment of  
CC protein A, here fused as a leader, and constituting a binding site for Fc  
CC The spacer contains the 11 C-terminal amino acids of the FB followed by  
CC Asp-Pro (a dilute acid cleavage site). The construct contains  
CC binding sites for both Fc and dioxin, and its structure can be  
CC summarised as: FB-Asp-Pro-VN-(Glyu-Ser)<sup>3</sup>-VH. The protein can be used for  
CC specific binding assays, affinity purification, biocatalysts, targeting,  
CC imaging an immunological treatment of oncogenic and infectious diseases.  
CC It offers further cleavage sites to circulating proteolytic enzymes, and  
CC improved stability. It reaches target tissues rapidly and is quickly  
CC cleared from the body. It also has reduced immunogenicity and its design  
CC facilitates binding to other moieties in drug targeting and imaging  
CC applications.  
QZ Sequence 251 AA;

Query Match	58.9%	Score 746:	DB 1:	length 251;
Best Local Similarity	58.5%	Pred No.	1.8e-44;	
Matches ~ 145;	Conservative	34;	Mismatches	57;
			Indels	12;
			Gaps	3

[illegible]

**RESULT 13**

15	W02191	standard; Protein; 239 AA.
16	W02101	

13-NOV-1996 (first entry)  
18-2-3/TRY202 single chain binding protein.  
Antibody engineering; monoclonal antibody; Mab;  
single chain antibody; immunoassay; fluorescein; 18-2-3/TRY202';  
single chain binding protein.  
Chimeric Mus musculus;  
Chimeric synthetic.  
US5518889-A.  
21-MAY-1996.  
21-MAY-1996.  
02-SEP-1986; 902971.  
02-SEP-1986; US-902971.  
02-SEP-1987; US-092110.  
19-JAN-1989; US-289617.  
25-APR-1990; US-512910.  
01-APR-1993; US-040440.  
06-JUN-1995; US-468988.  
(ENZO-) ENZON LABS INC.  
Bird RE, Hardman K, Ladner RC;  
WPI: 96-259060/26.  
N-PSDB; T36463.  
Immunoassay using single chain antigen binding mol. - as replacement  
for labelled or immobilised antibody, are less immunogenic, easier  
to engineer, more stable and less expensive  
Example 9; Fig 39A-B; 78pp; English.  
Computer-designed construct 18-2-3/TRY202' (W02191), encoded by  
the DNA sequence given in T36463, comprises the variable regions  
of the light and heavy chains of anti-fluorescein monoclonal  
antibody (Mab) 18-2-3 linked by a peptide designed to fit into a  
groove on the backside of the variable domain structure. The DNA  
construct was inserted into vector PGX3703 and introduced into E.  
coli. 18-2-3/TRY202' was produced as a single chain molecule. It  
exhibited biological binding activity equivalent in specifically  
and affinity to that of the original Mab.  
239 AA;

Query Match	58.8%	Score 745	DB 1	Length 239
Best Local Similarity	59.5%	Pred. No. 2e-44		
Matches 144	Conservative 40	Mismatches 50	Indels 8	Gaps 4

```

Oy 1 MDLDTGSPALLMSPGKGYMTGCRAPRPSV--SYMHMYXQOKGPGSPMPMYTTSNLSAGV 58
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MENVLTGSPALLMSPGKGYMTGCRASSVSSYTLHMYXQOKSGSPPLMYTGYHNSLASCV 60

Oy 59 PARPSGGSGSGTSTLYTSRYVEDAADAATYYCQOMSRSPTTEGGGSKLEIKSTGSGKSSSE 118
Db 61 PARPSGGSGSGTSTLYTSRYVEDAADAATYYCQOYSGYLTGAGAGKLEIKSSTGSG--SE 118

```



Search completed: May 13, 1999, 10:02:41  
Job time: 1499 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 10:03:48 ; Search time 21.53 Seconds  
(without alignments)  
417.579 Million cell updates/sec

Title: US-08-704-178-4

Perfect score: 1271  
Sequence: 1 MOLTOSPAIMASPGKRYTM.....SAYIANYGQGTSTVSSY 240

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database :

PIR\_58:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	560	44.1	287	3	PC4402	haptan-specific s1
2	526	41.4	137	2	S52445	Ig heavy chain V r
3	520.5	41.0	120	2	A34871	Ig kappa chain V r
4	496	39.0	107	2	PT0398	Ig light chain V r
5	492	38.7	107	2	PT0402	Ig light chain V r
6	491.5	38.7	235	2	S25058	Ig heavy chain - m
7	487.5	38.4	136	2	S04576	Ig heavy chain pre
8	486	38.2	107	2	PT0403	Ig light chain V r
9	482	37.9	107	2	PT0400	Ig light chain V r
10	482	37.9	107	2	PT0401	Ig light chain V r
11	480	37.8	99	2	S51210	Ig kappa chain V r
12	480	37.8	107	2	PT0397	Ig light chain V r
13	475	37.4	120	2	S03471	Ig heavy chain V-D
14	474	37.3	107	2	PT0385	Ig light chain V r
15	470	37.0	178	2	S29594	Ig gamma chain V r
16	469	36.9	107	2	A30562	Ig kappa chain V r
17	468	36.8	107	2	B30562	Ig kappa chain V r
18	464.5	36.5	122	2	S06823	Ig heavy chain V r
19	463	36.4	118	2	S25174	Ig heavy chain V r
20	461	36.3	130	1	UL0079	Ig kappa chain pre
21	460	36.2	107	2	PT0386	Ig light chain V r
22	460	36.2	117	2	S17586	Ig heavy chain V r
23	459	36.1	107	2	PT0399	Ig light chain V r
24	458.5	36.1	268	2	A56446	Ig heavy chain V r
25	458	36.0	103	2	S29591	Ig kappa chain V r
26	456.5	35.9	118	2	G37267	Ig heavy chain V r
27	456	35.9	106	2	G27887	Ig kappa chain V r
28	456	35.9	106	2	PS0071	Ig kappa chain V r
29	456	35.9	108	2	PL0278	Ig kappa chain V r
30	455.5	35.8	116	2	G30560	Ig kappa chain V r
31	455	35.8	104	2	S24289	Ig gamma chain V r
32	454	35.7	104	2	B49049	Ig light chain V r
33	453	35.6	130	2	A32513	Ig kappa chain pre
34	452	35.6	108	2	PL0276	Ig kappa chain V r
35	452	35.6	108	2	PL0277	Ig kappa chain V r
36	450	35.4	106	2	B54378	Ig light chain V r
37	445	35.0	130	2	B32456	Ig kappa chain pre
38	444	34.9	107	2	PC4405	anti-AFP antibody
39	443	34.9	107	2	PT0406	Ig light chain V r

## ALIGNMENTS

40	443	34.9	120	2	S03484	Ig heavy chain V-D
41	442	34.8	140	2	PL0013	Ig kappa chain pre
42	442	34.8	446	2	S40295	Ig kappa chain pre
43	442	34.8	246	2	S38950	Ig gamma-2a chain
44	441	34.7	107	2	S11119	Ig gamma chain - m
45	440	34.6	107	2	S11118	Ig kappa chain V r

## RESULT 1

PC4402

haptan-specific single-chain antibody variable fragment and alkaline phosphatase fusi





```

Query Match          37.8%   Score 480; DB 2; Length 99;
Best Local Similarity 93.8%; Fred. No. 3.2e-30;
Matches      90; Conservative    3; Mismatches     3; Indels       0; Gaps        0;

QY      3 LNPSPAINASGGEVYTMTCSSASSVSNNHWYQOKSSTSPKRLMYVDTSKLASGVPRSG 62
         |||||
DB       4 LNPSPAINASGGEVYTMTCSSASSYSMMHWIQOKSSISPKMLIDTISKLTISGVPRFSG 63
         |||||

QY      63 SGGNSYSTLTISSMEADATYYCYOGSGPYTFGS 98
         |||||
DB       64 SGGNSYSTLTISSMEADATYYCFQGGSCPYLTFGS 99
         |||||

RESULT 12
PRO397
IG light chain V region (SI07/VH11 group 2-10) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
C:Accession: PT03397
R:Behar, S.M.; Lustgarten, D.L.; Corbet, S.; Scharff, M.D.
J. Exp. Med. 173, 731-741, 1991
A:Title: Characterization of somatically mutated SI07 VH11-encoded anti-DNA autoantibody
A:Reference number: PT0336
A:Accession: PT03397
A:Molecule type: DNA
A:Residues: 1-107 <BEH>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

```

```

Best Local Similarity 80.0%: Pred. No. 9.6e-30:
Matches 96; Conservative 4; Mismatches 18; Indels 2; Gaps 1;

QY 120 VQLQSGVELYRGALVFLSCASPDENIKDYYIHMKQRPQGLGEMIGWHPENGNTYVD 179
      ||||| ||| : ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 VQLQSGALVELYRGASVFLSTAGSEFNKIDYIMHWKQRPQGLGEMIGRIDPANGNTKYD 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 180 PRFGKASITADTSSNAAVYLDLSLTSEDYNAVYICASYIYSATYATMYMGQGSIVYSS 239
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 PRFGKATITADTSSNTAYLDLSLTSEDYNAVYICARGMLRRD--AMDYMGQGSIVYSS 119
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
PR0395
Ig light chain V region (SI07/VH11 group 2-16) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
C:Accession: PR0395
R:Behar, S.M.: Lustgarten, D.L.: Corbet, S.: Scharrf, M.D.
J. Exp. Med. 173, 731-741, 1991
A:Title: Characterization of somatically mutated SI07 VH11-encoded anti-DNA autoantib
A:Reference number: PT0376
A:Accession: PR0395
A:Molecule type: DNA
A:Residues: 1-107 <BEH>
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: Immunoglobulin

```

Query Match 37.3%; Score 474; DB 2; Length 107;  
Best Local Similarity 85.4%; Pred. No. 1e-29;  
Matches 88; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Query Match 37.8%; Score 480; DB 2; Length 107;  
 Best Local Similarity 86.4%; Pred. No. 3.5e-30;  
 Matches 89; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```

QY      3 LHSQPIWASGEGEYITMTCSASSVSNMHWQOKSSTSPKRIYVDSTSLASGYGRSG 62
      |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      4 LHSPIWASGGRKVTYTCRSARSRSTYXNMITYQKSSISPKRIITIDISKLASGVFRSG 63
      |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
QY      63 SSGNSYSLTISMEADATYYCYOGSGYPPTFGSGTKLEIK 105
      |||.....|.....|.....|.....|.....|.....|.....|.....|.....|
DB      64 SSGNSYSLTISMEADATYYCFQSGSGYPLTFGSGTKLEIK 106
      |||.....|.....|.....|.....|.....|.....|.....|.....|.....|

```

QY	3	LTPSPALMAMSPPEKXTMTMCSSASSYSNMNHWYQOKSSRPLTWYDPLKSLASGPPGRFSG	62
Db	4	LTPSPALMAMSPPEKXTMTMCSSASSYSNMNHWYQOKSSRPLTWYDPLKSLASGPPGRFSG	63
QY	63	SGSGNSTSLTISMEADATATTCYCGSGVPTFGSGTKLEIK	105
Db	64	SGSGNSTSLTISMEADATATTCYCGSGVPTFGSGTKLEIK	106

RESULT 15  
29594

RESULT	Ig gamma chain (MM65) - mouse (fragment)
13	C:Species: Mus musculus (house mouse)
S03471	C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
Ig heavy chain V-D-J region (hybridoma G5 Bb 2.2) - mouse (fragment)	C:Accession: S29594

submitted to the EMBL Data Library, February 1991  
:Reference number: S29593

R.Rocca-Serra J.; Matthes, H.W.; Kaartinen, M.; Milstein, C.; Theze, J.; Fougereau, M.  
 A: J. 2, 867-872, 1983  
 A:Title: Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT  
 hypervariable regions.  
 A:Reference number: S03471; MUID:84057768  
 A:Accession: S03471

A:Cross-references: EMBL:X57857; NID:g52590; PID:g52591  
C:Keywords: Immunoglobulin

A: Molecule type: mRNA  
A: Residues: 7-120 <R0C1>  
A: Cross-References: EMBL:X01820; NID:G51833; PID:G29256; PID:G1333983  
A: Note: this sequence was determined from the differentiated gene  
R: Rocca-Serra, J.; Malek, J. C.; Molnár, D.; Leclercq, L.; Somme, G.; Theze, J.; Fougere, J. Immunol. 129, 2554-2558, 1982  
A: Title: The limited diversity of the mouse gamma-chains anti-GAT repertoire does not se  
A: Reference number: S07453; MUID:83058021

Query Match	37.0%	Score 470;	DB 2;	Length 178;
Best Local Similarity	74.4%	Pred. No. 3.5e-29;		
Matches 96;	Conservative 7;	Mismatches 20;	Indels 6;	Gaps 2;
QY	111	SKSSBEGCVQLQSGVELREGALVKYISCKASDINIDYTHHWKQRPDGLNIGWTH	170	
:	:	:	:	:

A:Accession: S07453  
A:Molecule type: protein  
A:Residues: 1-43 <R0C2>  
C:Superfamily: Immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin

```
Qy      171 PGNNTVYDPKFEQGKASITADTSSNAVLQLSSLTSEPTAVYYCASYYYYSAFYAMYWG 230
        | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      66 PANGTEYDPEFQGKAITADSTINAYLQLSSLTSEDPAVYYCTCGN---AYGDYWG 122
```

Query Match 37.48; Score 475; DB 2; Length 120;

QY 231 QGTSVTVSS 239  
|||||

Thu May 13 10:28:22 1999

us-08-704-178-4.rpt

Page 5

Db 123 QGTSVTWSS 131

Search completed: May 13, 1999, 10:03:48  
Job time: 1380 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 10:05:48 ; Search time 16.53 Seconds

(without alignments)  
389,696 Million cell updates/sec

Title: US-08-704-178-4

Perfect score: 1271

Sequence: 1 MQLTSPALMSASPGKVTM.....SAYMYTNGGTSTVSSY 240

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	450	35.4	107	1 KV6F_MOUSE	P04940 mus musculu
2	446	35.1	107	1 KV6J_MOUSE	P04944 mus musculu
3	445	35.0	107	1 KV6I_MOUSE	P04943 mus musculu
4	444	34.9	107	1 KV6G_MOUSE	P04941 mus musculu
5	444	34.9	107	1 KV6H_MOUSE	P04942 mus musculu
6	412	32.4	107	1 KV6I_MOUSE	P01676 mus musculu
7	410	32.3	108	1 KV6K_MOUSE	P04945 mus musculu
8	409	32.2	107	1 KV6A_MOUSE	P01678 mus musculu
9	408	32.1	107	1 KV6D_MOUSE	P01677 mus musculu
10	407	32.0	107	1 KV6C_MOUSE	P01677 mus musculu
11	406	31.9	120	1 HV03_MOUSE	P01747 mus musculu
12	401.5	31.6	140	1 HV02_MOUSE	P01746 mus musculu
13	401	31.5	107	1 KV6J_MOUSE	P01791 mus musculu
14	400	31.5	139	1 HV07_MOUSE	P01759 mus musculu
15	395	31.1	129	1 KV4A_MOUSE	P01680 mus musculu
16	370	29.1	117	1 HV13_MOUSE	P01757 mus musculu
17	367.5	28.9	118	1 HV51_MOUSE	P06330 mus musculu
18	366	28.8	121	1 HV01_MOUSE	P01745 mus musculu
19	365	28.7	117	1 HV12_MOUSE	P01756 mus musculu
20	363	28.6	137	1 HV11_MOUSE	P01755 mus musculu
21	361.5	28.4	120	1 HV50_MOUSE	P06329 mus musculu
22	361	28.4	138	1 HV48_MOUSE	P03980 mus musculu
23	356	28.0	117	1 HV52_MOUSE	P06327 mus musculu
24	352.5	27.7	143	1 HV1C_HUMAN	P01744 homo sapien
25	346	27.2	117	1 HV06_MOUSE	P01750 mus musculu
26	340	26.8	136	1 HV15_MOUSE	P01759 mus musculu
27	337.5	26.6	117	1 HV09_MOUSE	P01753 mus musculu
28	336	26.4	117	1 HV04_MOUSE	P01748 mus musculu
29	330	26.0	117	1 HV14_MOUSE	P01758 mus musculu
30	328.5	25.8	108	1 KV1R_HUMAN	P01610 homo sapien
31	326.5	25.7	108	1 KV5K_MOUSE	P01640 mus musculu
32	326	25.6	117	1 HV49_MOUSE	P06332 mus musculu
33	325.5	25.6	108	1 KV1B_HUMAN	P01594 homo sapien
34	325.5	25.6	108	1 KV1V_HUMAN	P04430 homo sapien
35	325	25.6	117	1 HV05_MOUSE	P01749 mus musculu
36	324.5	25.5	108	1 KV5O_MOUSE	P01648 mus musculu
37	323.5	25.5	108	1 KV5L_MOUSE	P01645 mus musculu
38	322.5	25.4	108	1 KV5M_MOUSE	P01645 mus musculu
39	322.5	25.4	108	1 KV1P_HUMAN	P01608 homo sapien
40	322.5	25.3	129	1 KV1M_HUMAN	P04431 homo sapien
41	322	25.3	109	1 KV3F_HUMAN	P01622 homo sapien
42	321.5	25.3	108	1 KV5D_MOUSE	P01635 mus musculu
43	321	25.3	117	1 HV10_MOUSE	P01754 mus musculu

## ALIGNMENTS

44 320.5 25.2 108 1 KV1L\_HUMAN P01604 homo sapien  
45 320.5 25.2 108 1 KV5N\_MOUSE P01647 mus musculu

## RESULT 1

ID KV6F\_MOUSE STANDARD; PRT; 107 AA.

AC P04940;  
DT 13-AUG-1987 (REL. 05, CREATED)  
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE IG KAPPA CHAIN V-VI REGION (NO2-17.4.1.).  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE; 83271467.  
RX KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;  
RL NATURE 304:320-324(1983).  
CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
DR EMBL; K00735; G196435; -.  
DR HSSP; P01679; 1FVB.  
KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.  
FT DOMAIN 1 23 FRAMEWORK 1.  
FT DOMAIN 24 33 COMPLEMENTARITY-DETERMINING 1.  
FT DOMAIN 34 48 FRAMEWORK 2.  
FT DOMAIN 49 55 COMPLEMENTARITY-DETERMINING 2.  
FT DOMAIN 56 87 FRAMEWORK 3.  
FT DOMAIN 88 96 COMPLEMENTARITY-DETERMINING 3.  
FT DISULFID 97 106 FRAMEWORK 4.  
FT NON\_TER 23 87 BY SIMILARITY.  
SQ SEQUENCE 107 AA; 11561 MW; 1BCOC1E3 CRC32;

Query Match Best Local Similarity 35.4%; Score 450; DB 1; Length 107;  
Matches 86; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 3 LTSPALMSASPGKVTMTCSASSVSNMAYOQKSTSPKLMWYDTSKLASGVPRFSG 62  
Db 4 LTSPALMSASPGKVTMTCSASSVSNMAYOQKSTSPKLMWYDTSKLASGVPRFSG 63  
QY 63 SSGNSSTSLTSSKAEADATYYCYGSGYPTFGSTKLEIK 105  
Db 64 SSGATSYSLTSSKQADADATYYCYGSGYPTFGSTKLEIK 106

RESULT 2  
ID KV6J\_MOUSE STANDARD; PRT; 107 AA.

AC P04944;  
DT 13-AUG-1987 (REL. 05, CREATED)  
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE IG KAPPA CHAIN V-VI REGION (NO5-78.2.6).  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE; 83271467.  
RX KAARTINEN M., GRIFFITHS G.M., MARKHAM A.F., MILSTEIN C.;  
RL NATURE 304:320-324(1983).  
CC -1- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.  
DR EMBL; K00744; G196453; -.  
DR HSSP; P01679; 1FVB.  
KW IMMUNOGLOBULIN V REGION; HYBRIDOMA.  
FT DOMAIN 1 23 FRAMEWORK 1.

Query Match	35.0%:	Score 445:	DB 1:	Length 107;	
Best Local Similarity	82.5%:	Pred. No. 5.1e-28:			
Matches	85;	Conservative	6;	Mismatches 12;	Indels 0;
				Gaps	0.
QY	3	LTGSPAIMSAPGGEKVTMTCSASSSVSNMHWYQAKSTSPKLMVYDTSKLASGVGRFSG	62		
Db	4	LTGSPAIMSAPGKQVNTMTCSASSSVSNMHWYQAKSGTSPKRMVYDTSKLASGXPARFSG	63		
QY	63	SGSGNSYSLTITSSMEADPAITYCTCQSGSGYPTTFESSGTFLEIK	105		
Db	64	SGSATSYSLTITSSMOAEDPAITYCCOOWSSNPLTFAGTKLEIK	106		
RESULT	4				
KV6G_MOUSE					
IQ_KV6G_MOUSE		STANDARD:	PRT:	107 AA.	
IC_P045M1:					

RESULT	5				
KV6H_MOUSE					
ID	KV6H_MOUSE	STANDARD;	PRT;	107	AA.
AC	P04942;				
DT	13-AUG-1987 (REL. 05, CREATED)				
DT	13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)				
DT	01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)				
DE	IG KAPPA CHAIN V-VI REGION (MOS-61.1.2).				
OS	MOS MUSCULUS (MOUSE).				
OC	EBKAROTIA; METAPOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 83271467.				
RA	KARLITENEN M., GRIEFITHS G.M., MARKHAM A.F., MILSTEIN C.;				
RL	NATURE 304:320-324(1983).				
CC	-I- FUNCTION: ANTI-2-PHENYL OXAZOLONE (PHOX) ANTIBODY.				
DR	EMBL; K00739; G196443; -.				
DR	HSP; P01679; 1FV.				
KW	IMMUNOGLOBULIN V REGION; HYBRIDOMA.				
FT	DOMAIN	1	23		
FT	DOMAIN	24	33		
FT	DOMAIN	34	48		
FT	DOMAIN	49	55		
FT	DOMAIN	56	87		
FT	DOMAIN	88	96		
FT	DOMAIN	97	106		
FT	DISULFID	23	87		
FT	NON_TER	107	107		
QO	SEQUENCE	107	AA;	11605	MM; 1CF3CEBD CRC32;





```

RESULT 9
KV6D_MOUSE STANDARD; PRT; 107 AA.
ID KV6D_MOUSE
AC P01678:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE IG KAPPA CHAIN V-VI REGION (SAPC 10).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP MEDLINE: 81054757.
RA RUDIKOFF S., RAO D.N., GLAUDEMANS C.P.J., POTTER M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 77:4270-4274(1980).
CC -1- THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT BIND GALACTAN.
DR HSP; P01679; ZFBJ.
KW IMMUNOGLOBULIN V REGION.
FT DOMAIN 1 23 FRAMEWORK 1.
FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 3 48 FRAMEWORK 2.
FT DOMAIN 4 55 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 5 56 FRAMEWORK 3.
FT DOMAIN 6 87 COMPLEMENTARITY-DETERMINING 3.
FT DOMAIN 7 96 FRAMEWORK 4.
FT DISULFID 23 87 BY SIMILARITY.
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11554 MW; 609EDE13 CRC32;

Query Match 32.1%; Score 408; DB 1; Length 107;
Best Local Similarity 79.8%; Pred. No. 3.5e-25;
Matches 83; Conservative 6; Mismatches 13; Indels 2; Gaps 2;

QY 3 LTOSPAINASPGKVTMTCSASSSVSNMHWYQOKSSTPKLWYDTSKLASGVGPRFSG 62
DB 4 LTOSPAINASLGQKVITTCSSASSSVSYMHYQOKSGTSPKWIYEISKLASGVPRFSG 63
QY 63 SSGNSYSLTSSMEDIADATYYCYGSGYP-FTFGSGTKLEIK 105
DB 64 SSGSTSYSLTSSMEDIADATYYCOQWN-YPLTFGGTKLEIK 106

RESULT 10
KV6C_MOUSE STANDARD; PRT; 107 AA.
ID KV6C_MOUSE
AC P01677:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE IG KAPPA CHAIN V-VI REGION (TEPC 601 AND TEPC 191).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP MEDLINE: 79082830.
RA RAO D.N., RUDIKOFF S., POTTER M.;
RL BIOCHEMISTRY 17:5555-5559(1978).
RN [2]
RP MEDLINE: 81054757.
RA RUDIKOFF S., RAO D.N., GLAUDEMANS C.P.J., POTTER M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 77:4270-4274(1980).
CC -1- THE TWO SEQUENCES ARE IDENTICAL.
CC -1- THESE CHAINS WERE ISOLATED FROM MYELOMA PROTEINS THAT BIND GALACTAN.
DR HSP; A01941; KVMX4.
DR HSP; P01679; ZFBJ.
KW IMMUNOGLOBULIN V REGION.

```

```

FT DOMAIN 1 23 FRAMEWORK 1.
FT DOMAIN 2 33 COMPLEMENTARITY-DETERMINING 1.
FT DOMAIN 3 48 FRAMEWORK 2.
FT DOMAIN 4 55 COMPLEMENTARITY-DETERMINING 2.
FT DOMAIN 5 56 FRAMEWORK 3.
FT DOMAIN 6 87 COMPLEMENTARITY-DETERMINING 3.
FT DOMAIN 7 96 FRAMEWORK 4.
FT DISULFID 23 87 BY SIMILARITY.
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11568 MW; 154F6572 CRC32;

Query Match 32.0%; Score 407; DB 1; Length 107;
Best Local Similarity 78.8%; Pred. No. 4.2e-25;
Matches 82; Conservative 8; Mismatches 12; Indels 2; Gaps 2;

QY 3 LTOSPAINASPGKVTMTCSASSSVSNMHWYQOKSSTPKLWYDTSKLASGVGPRFSG 62
DB 4 LTOSPAINASLGQKVITTCSSASSSVSYMHYQOKSGTSPKWIYEISKLASGVPRFSG 63
QY 63 SSGNSYSLTSSMEDIADATYYCYGSGYP-FTFGSGTKLEIK 105
DB 64 SSGSTSYSLTSSMEDIADATYYCOQWN-YPLTFGGTKLEIK 106

RESULT 11
KV03_MOUSE STANDARD; PRT; 120 AA.
ID KV03_MOUSE
AC P01747:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 21-JUL-1986 (REL. 01, LAST ANNOTATION UPDATE)
DE IG HEAVY CHAIN V REGION (36-65).
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP MEDLINE: 83131846.
RA STEKELYITZ M., GEFTER M.L., BRODEUR P., RIBLET R.,
RA MARSHAK-ROTHSTEIN A.;
RL EUR. J. IMMUNOL. 12:1023-1032(1982).
CC -1- FROM ANALYSIS OF THE SIZES OF SEVERAL OTHER DIFFERENTIATED GENES
CC THAT HYBRIDIZE TO THIS ONE, THE AUTHORS CONCLUDE THAT ALL OF
CC THESE V REGIONS HAVE REARRANGED TO THE SAME J SEGMENT, JH2.
DR HSP; A02028; HYMSC7.
DR HSP; P01789; GFAB.
KW IMMUNOGLOBULIN V REGION; ANTIBODIES; ANTIBODY; HYBRIDOMA.
FT NON_TER 120 120
SQ SEQUENCE 120 AA; 13307 MW; BBABCA1 CRC32;

Query Match 31.9%; Score 406; DB 1; Length 120;
Best Local Similarity 63.3%; Pred. No. 5.7e-25;
Matches 76; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

QY 120 VOLQOSGVELVRCGALVKLSCKASDENIKDYIHWYQOREGQLEWIGMHPENGNTVYD 179
DB 1 VOLQOSGAEIVRAGSSVYKMSCKASGYFTSTYGINWYQOREGQLEWIGMHPNGYTKYN 60
QY 180 PKFGKASITADTSSNAAYQLSLTSEDTAVVYCAASYSAAYANYWYGOGSTVYSS 239
DB 61 EKFGKTTLVNDKSSSTAYWQLRSLTSEDSAVYFCASVYGGSYDYNGOGTTLVSS 120

RESULT 12
KV02_MOUSE STANDARD; PRT; 140 AA.
ID KV02_MOUSE
AC P01746:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE IG HEAVY CHAIN PRECURSOR V REGION (9367).

```





GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 10:05:19 ; Search time 28.48 Seconds

(without alignments)  
464,909 Million cell updates/sec

Title: US-08-704-178-4

Perfect score: 1271

Sequence: 1 MGLTSPALMSASPEKVTM.....SAYIANYWGQSTVTSY 240

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database :

SPTREMBL.8.\*  
1: sp\_fungi.\*  
2: sp\_human.\*  
3: sp\_invertebrate.\*  
4: sp\_mammal.\*  
5: sp\_mhc.\*  
6: sp\_organelle.\*  
7: sp\_phage.\*  
8: sp\_plant.\*  
9: sp\_bacteria.\*  
10: sp\_rodent.\*  
11: sp\_virus.\*  
12: sp\_vertebrate.\*  
13: sp\_unclassified.\*  
14: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	424	33.4	130	10	P80913	P80913 mus musculi
2	212	16.7	100	2	O15535	O15535 homo sapien
3	188	14.8	85	2	O75724	O75724 homo sapien
4	182	14.3	77	2	O75741	O75741 homo sapien
5	181.5	14.3	81	2	O75736	O75736 homo sapien
6	181	14.2	81	2	O75734	O75734 homo sapien
7	176	13.8	77	2	O75726	O75726 homo sapien
8	175.5	13.8	75	2	O75743	O75743 homo sapien
9	174	13.7	82	2	O75725	O75725 homo sapien
10	174	13.7	80	2	O75727	O75727 homo sapien
11	174	13.7	77	2	O75728	O75728 homo sapien
12	171	13.5	72	2	O75738	O75738 homo sapien
13	168	13.2	76	2	O75733	O75733 homo sapien
14	167	13.1	86	2	O75740	O75740 homo sapien
15	166	13.1	78	2	O75739	O75739 homo sapien
16	163	12.8	81	2	O75719	O75719 homo sapien
17	163	12.8	81	2	O75721	O75721 homo sapien
18	162.5	12.8	82	2	O75732	O75732 homo sapien
19	162	12.7	88	2	O75737	O75737 homo sapien
20	162	12.7	76	2	O75742	O75742 homo sapien
21	160.5	12.6	82	2	O75729	O75729 homo sapien
22	158	12.4	78	2	O75730	O75730 homo sapien
23	158	12.4	79	2	O75731	O75731 homo sapien
24	157	12.4	80	2	O75735	O75735 homo sapien
25	155	12.2	74	2	O75744	O75744 homo sapien
26	150	11.8	78	2	O75723	O75723 homo sapien
27	147	11.6	86	2	O75722	O75722 homo sapien
28	140.5	11.1	97	2	O43234	O43234 homo sapien
29	135.5	10.7	137	5	O31181	O31181 mus musculi

30	135	10.6	509	10	P97710	P97710 rattus norv
31	134	10.5	258	12	O90556	O90556 ginglymosto
32	134	10.5	250	12	O90569	O90569 ginglymosto
33	133.5	10.5	398	2	O00241	O00241 homo sapien
34	133	10.5	158	12	O90531	O90531 ginglymosto
35	132	10.4	78	2	O75720	O75720 homo sapien
36	131.5	10.3	50	2	O15533	O15533 homo sapien
37	131.5	10.3	503	2	P78324	P78324 homo sapien
38	131.5	10.3	258	12	O90559	O90559 ginglymosto
39	130.5	10.3	145	2	O16237	O16237 homo sapien
40	130	10.2	418	10	O70426	O70426 rattus norv
41	129.5	10.2	237	12	O90545	O90545 ginglymosto
42	129.5	10.2	254	12	O90557	O90557 ginglymosto
43	127	10.0	133	5	O31178	O31178 mus musculi
44	125.5	9.9	509	10	O08907	O08907 mus musculi
45	125.5	9.9	509	10	O35924	O35924 mus musculi

## ALIGNMENTS

RESULT 1  
P80913 PRELIMINARY; PRT; 130 AA.

AC P80913;  
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE IG KAPPA CHAIN V REGION PRECURSOR.  
OC MUS MUSCULUS (MOUSE)  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCITUROGNATHI; MORIDAE; MORINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA JANON K.;  
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL, X79906; E269393; ALT-TERM.  
KW IMMUNOGLOBULIN V REGION; SIGNAL.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 130 IG KAPPA CHAIN V REGION.  
SQ SEQUENCE 130 AA; 14124 MW; 1DC6F7D9 CRC32;

Query Match 33.4%; Score 424; DB 10; Length 130;  
Best Local Similarity 79.0%; Pred. No. 9.6e-26;  
Matches 83; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

QY 3 LTSPALMSASPEKVTMTCSASSVSN--MHWIOKSSTPKIMVYDISKLASGVGRF 60  
|||||  
Db 26 LTSPALMSASLGERVTMTCTASSSVSSSYLHWYQKFGSSPKIMVYSTSLASGVPARF 85  
61 SSGSGNSYSLTLSMEAEADAATVYCYGSGPYETFGSGKLEIK 105  
Db 86 SSGSGTSTSLTSSMEAEADAATVYCHQYHFPHTFGGKLEIK 130

RESULT 2  
O15535 PRELIMINARY; PRT; 100 AA.

AC O15535;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE V KAPPA (FRAGMENT).  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA KATO S., TACHIBANA K., TAKAYAMA N., KATAOKA H., YOSHIDA M.C.,  
RA TAKANO T.;  
RL SUBMITTED (SEP-1990) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL, D90161; G1262585; -



RP	SEQUENCE FROM N.A.
RC	TISSUE-INTESTINE;
RA	FISCHER M., KUEPPERS R.;
RT	"Human IgA and IgM secreting intestinal plasma cells carry heavily
RT	mutated VH region genes.";
RL	SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR	EMBL: AJ009531; E1311462; -.
FT	NON_TER
FT	1
FT	81
FT	81
SEQ	SEQUENCE 81 AA; 9040 MW; 56BA902 CRC32;

Query Match	14.2%;	Score 181;	DB 2;	length 81;
Best Local Similarity	42.5%;	Pred. No. 1.9e-07;		
Matches 34;	Conservative 13;	Mismatches 33;	Indels 0;	Gaps 0;

```

Oy 150 YYHHVKKRPPROPGLEWIGIMHEPKNVYDPDKPGKASITADTSNNAAYLOLSITSEPT 208
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2 YSMNNVROAPKPGKGLEWVSSISSTSSNIFDGVAOGRFAISHDNAKNSLTIQISLGADPT 61

Oy 210 AVYYCASYYYYSAVYAMYYW 229
      |||||
Db 62 AVYYCARLNTKRYENGMDYW 81

```

RESULT 7  
075726  
ID 075726 PRELIMINARY; PRT; 77 AA.

DT 01-NOV-1998 (TREMBL.REL. 08, CREATED)  
DT 01-NOV-1998 (TREMBL.REL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBL.REL. 08, LAST ANNOTATION UPDATE)  
DE 1G HEAVY CHAIN VARIABLE REGION (FRAGMENT).  
GN VH.  
OS HOMO SAPIENS (HUMAN).  
OS EUCARYOTA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
OC [1]  
RN RP SEQUENCE FROM N.A.  
RC TISSUE-INTESTINE;  
RC FISCHER M., KUEPPERS R.;  
RA "Human IgA and IgM secreting intestinal plasma cells carry heavily  
RT mutated VH region genes."  
RT SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
RL EMBL, AJ009523; E1311446; -.  
DR NON\_TER 1  
FT NON\_TER 77  
FT SEQUENCE 77 AA; 9022 MW; EB7B458D CRC32;  
SQ

Query Match	13.8%	Score 176;	DB 2;	Length 77;
Best Local Similarity	42.9%	Pred. No. 4.4e-07;		
Matches 30; Conservative	17;	Mismatches 23;	Indels 0;	Gaps 0;

[illegible]

RESULT	8	
075743		
ID	075743	PRELIMINARY; PRT; 75 AA

DT 01-NOV-1998 (TREMBLREL, 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL, 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)  
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).  
CN VH.  
OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA,  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

QY 151 YIHWKQRPQEGLEWITGWIHPENGNTVYDPKFGKASITADTSSNAAYLQISLTSEDTA 210

Db 1 YMAWTRQPPGKGLEWITG-G-HYYPGGTYNPSLSRVITISADTSKHFFSLKMTSVTADTA 59

RESULT	9	
075725		
ID	075725	PRELIMINARY;
		PRT;
		82 AA

DT 01-NOV-1998 (TREMBLERL. 08, CREATED)  
 DT 01-NOV-1998 (TREMBLERL. 08, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLERL. 08, LAST ANNOTATION UPDATE)  
 DE VH HEAVY CHAIN VARIABLE REGION (FRAGMENT).  
 GN .  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-INTESTINE;  
 RA FISCHER M., KUEPERS R.;  
 RT "Human Iga and Igm secreting intestinal plasma cells carry heavily  
 RT mutated VH region genes."  
 RL EMBL (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AJ009522; E1311444; -.  
 FT NON\_TER 1  
 FT NON\_TER 82  
 SQ SEQUENCE 82 AA; 9566 MW; 1F2E1379 CRC32;

Query Match	13.7%;	Score 174;	DB 2;	Length 82;
Best Local Similarity	40.7%;	Pred. No. 6.8e-07;		
Matches	33;	Conservative	18;	Mismatches 28;
				Indels 2;
				Gaps 2;

[illegible]

RESULT 10  
075727  
ID 075727 PRELIMINARY; PRT; 80 AA

DT	01-NOV-1998	(TREMBLREL. 08, CREATED)
DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)

```

DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE.
RA FISCHER M., KUEPPERS R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RL mutated VH region genes.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AJ009524; E1311448; -.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 9351 MW; 63DE158A CRC32;

Query Match 13.7%; Score 174; DB 2; Length 80;
Best Local Similarity 44.7%; Pred. NO. 6.6e-07;
Matches 34; Conservative 16; Mismatches 24; Indels 2; Gaps 1;

QY 151 YIHWKORPEOGLIEWIGIHPENGN--TVYDPKFGKASITADTSSNAAYLQLSITSSED 208
Db 1 YMDWVRQAPGKGLWVGRKRKANSYTTETAYASVAGRFTISDDSKNSLYLQMSLKIDD 60
QY 209 TAVYYCASYYYSAAY 224
Db 61 TAVYYCARDHLRSAAY 76

RESULT 11
ID 075728 PRELIMINARY; PRT; 77 AA.
AC 075728:
DT 01-NOV-1998 (TREMBLREL. 08. CREATED)
DT 01-NOV-1998 (TREMBLREL. 08. LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08. LAST ANNOTATION UPDATE)
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE.
RA FISCHER M., KUEPPERS R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RL mutated VH region genes.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AJ009525; E1311450; -.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8925 MW; EALACFDF CRC32;

Query Match 13.7%; Score 174; DB 2; Length 77;
Best Local Similarity 43.9%; Pred. NO. 6.3e-07;
Matches 29; Conservative 15; Mismatches 22; Indels 0; Gaps 0;

QY 150 YIHWKORPEOGLIEWIGIHPENGNTVYDPKFGKASITADTSSNAAYLQLSITSSED 209
Db 1 YMNWIRQAPGEGIQWVSTVENGVTVYPPDSVKGRFSTSRDNSNTLYLQNSLRADBT 60
QY 210 AVYYCA 215
Db 61 ALYYCA 66

RESULT 12
ID 075738 PRELIMINARY; PRT; 72 AA.
AC 075738:
DT 075738:

```

```

DE 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN (1)
RN RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE.
RA FISCHER M., KUEPPERS R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RL mutated VH region genes.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AJ009537; E1311470; -.
FT NON_TER 1 1
FT NON_TER 72 72
SO SEQUENCE 72 AA; 8345 MW; E0FB044A CRC32;

Query Match 13.5%; Score 171; DB 2; Length 72;
Best Local Similarity 39.7%; Pred. No. 9.9e-07;
Matches 31; Conservative 16; Mismatches 23; Indels 8; Gaps 1;

OY 152 IHWAQREPGLEWIGWHPENGNTVDPKFGKASITADTSSNAAYLOLSLTSEDY 211
Db 3 MMWVQARQKGLWISYIRSGFTFYIDSVGRFTISSDIKNSVYLQMSLRDDTAV 62
OY 212 YYCASYYYSAAYMYW 229
Db 63 YYCAR-----DLAYW 72

RESULT 13
O75733 PRELIMINARY; PRT; 76 AA.
AC 075733;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).
GN VH.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN (1)
RN RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE.
RA FISCHER M., KUEPPERS R.;
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily
RL mutated VH region genes.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AJ009530; E1311460; -.
FT NON_TER 1 1
FT NON_TER 76 76
SO SEQUENCE 76 AA; 8543 MW; C95EB66A CRC32;

Query Match 13.2%; Score 168; DB 2; Length 76;
Best Local Similarity 43.8%; Pred. No. 1.8e-06;
Matches 35; Conservative 12; Mismatches 25; Indels 8; Gaps 2;

OY 152 IHWAQREPGLEWIGWHPENGNTVDPKFGKASITADTSSNAAYLOLSLTSEDY 209
Db 3 IHWAQRTSGKGLWVVARIRSKTSHATAYAAVSVEGRFYVSRDSSKNTAYLOMNSLKSEDY 62
OY 210 AYYCASYYYSAAYMYW 229
Db 63 AYYCA-----AGSSSEYW 76

RESULT 14

```



Search completed: May 13, 1999, 10:05:19  
Job time: 657 sec

075740  
ID 075740 PRELIMINARY; PRT; 86 AA.  
AC 075740;  
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).  
GN VH.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-INTESTINE;  
RA FISCHER M., KUEPPERS R.;  
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily  
mutated VH region genes."  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AJ009539; E1311474; -.  
FT NON\_TER 1  
FT NON\_TER 86  
SQ SEQUENCE 86 AA; 9625 MW; 076278BC CRC32;

Query Match 13.1%; Score 167; DB 2; Length 86;  
Best Local Similarity 45.6%; Pred. No. 2.4e-06;  
Matches 31; Conservative 15; Mismatches 20; Indels 2; Gaps 1;

QY 151 YIHWKQPEQGLEWIGMI--HPENGNTVYDPKFGKASITADTSSNAAYIQLSITSED 208  
DB 3 WMNVKQAPGKGLWVGKSKTDGTTDYAPVAKRFTISRDSKNTLYLQNSLAKTED 62

QY 209 TAVYCCAS 216  
DB 63 TAVYCCIT 70

RESULT 15  
075739  
ID 075739 PRELIMINARY; PRT; 78 AA.  
AC 075739;  
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE IG HEAVY CHAIN VARIABLE REGION (FRAGMENT).  
GN VH.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-INTESTINE;  
RA FISCHER M., KUEPPERS R.;  
RT "Human Iga and Igm secreting intestinal plasma cells carry heavily  
mutated VH region genes."  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AJ009538; E1311472; -.  
FT NON\_TER 1  
FT NON\_TER 78  
SQ SEQUENCE 78 AA; 9031 MW; 05D273B9 CRC32;

Query Match 13.1%; Score 166; DB 2; Length 78;  
Best Local Similarity 40.5%; Pred. No. 2.6e-06;  
Matches 32; Conservative 20; Mismatches 23; Indels 4; Gaps 2;

QY 151 YIHWKQPEQGLEWIGMIHPENGNTVYDPKFGKASITADTSSNAAYIQLSITSED 210  
DB 1 YMSWTRQSPGKGLWIGIYI-YTGSINNPKSKRVITSLGISENQFSLNLTSTVADTA 59

QY 211 VYCASTYYSAYIAYMYW 229  
DB 60 VYFCAR--HSSSWPIYF 75



GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 13, 1999, 10:02:41 ; Search time 23.53 Seconds  
(without alignments)  
206.293 Million cell updates/sec

Title: US-08-704-178-4

Sequence: 1 MQLTQSPAIMSASPEKVTM.....SAYYAMYWGQGTSTVSSY 240

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A\_Geneseq\_34:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1269	99.4	239	1	W15186	Single-chain anti-
2	1254	99.1	238	1	R45443	Sequence of the s
3	829.5	65.3	244	1	R60780	Fv(Gp-2) immuno
4	821.5	64.6	241	1	R12788	OVB3 light and hee
5	795.5	62.6	237	1	R94020	Anti-erbB2 scFv. I
6	795.5	62.6	237	1	W15185	Single-chain anti-
7	790.5	62.2	236	1	R45442	Sequence of the s
8	763	60.0	240	1	R15443	Single chain Fv fr
9	759	59.7	240	1	W25784	Anti-gp130 antibod
10	752	59.2	553	1	W11508	Single chain anti-
11	752	59.2	553	1	W73223	Single chain anti-
12	751.5	59.1	239	1	W02181	H22 anti-CERA anti
13	748.5	58.9	239	1	R43679	18-2-3/TRV202' sin
14	748.5	58.9	239	1	R99646	Single chain poly
15	744.5	58.6	241	1	R06482	Single chain bindin
16	727.5	57.2	245	1	W25783	18-2-3-/TRV202' . S
17	725.5	57.1	473	1	W26646	Anti-T3 antibody c
18	725.5	57.1	514	1	W26647	Chimeric receptor
19	725.5	57.1	403	1	W26648	Chimeric receptor
20	725.5	57.1	651	1	W26649	Chimeric receptor
21	725.5	57.1	692	1	W26650	Chimeric receptor
22	725.5	57.1	382	1	W26651	Chimeric receptor
23	724	57.0	250	1	R43678	Single chain poly
24	724	57.0	250	1	R99648	Single chain bindin
25	724	57.0	250	1	W02190	Single chain poly
26	719	56.6	246	1	R43677	Single chain poly
27	715	56.3	250	1	R06481	Single chain poly
28	715	56.3	246	1	R99647	Single chain bindin
29	714	56.2	486	1	R37649	Sequence of the s
30	712.5	56.1	271	1	R20185	Sequence of the s
31	712	56.0	262	1	R37380	Sequence of the s
32	712	56.0	264	1	R97381	PLAP CC49/212 SCA
33	711	55.9	242	1	R33680	Single chain poly
34	711	55.9	242	1	R99650	Single chain bindin
35	711	55.9	242	1	W02192	18-2-3-/TRV59 sing
36	701	55.2	242	1	R06483	18-2-3-/TRV59. Sin
37	700.5	55.1	360	1	W35128	R. p18piens recomb
38	700.5	55.1	358	1	W35127	R. p18piens recomb
39	700.5	55.1	358	1	W35130	R. p18piens recomb
40	699	55.0	246	1	W02189	TRV59 single chain
41	698.5	55.0	379	1	W35126	R. p18piens recomb
42	698.5	55.0	355	1	W35129	R. p18piens recomb
43	698.5	55.0	355	1	W35125	R. p18piens recomb

## ALIGNMENTS

44	696	54.8	301	1	W37085	Anti-human SC sing
45	691	54.4	243	1	R60781	Fv(GP-4) immunosup
ALIGNMENTS						
RESULT	1					
ID	W15186	W15186 standard; Protein; 239 AA.				
AC	W15186;					
DT	05-JUN-1997	(first entry)				
DE	Single-chain anti-erbB2 antibody e21 (Fv).					
KW	Single chain antibody; variable region; light chain; heavy chain;					
RV	breast cancer; ovarian cancer; non-small cell lung carcinoma;					
OS	Immunodiagnosis; treatment; cytotoxic agent; erbB-2.					
MS	Mus musculus.					
OS	Synthetic.					
FT	Key	location/Qualifiers				
FT	region	106..119				
FT		/label= linker				
PN	US5587458-A.					
PD	24-DEC-1996.					
PF	07-OCT-1991; 772270.					
PR	07-OCT-1991; US-772270.					
PR	30-JUN-1992; US-906555.					
PR	14-MAY-1993; US-061092.					
PA	(ARON-) ARONEX PHARM INC.					
PI	Bird re, Kasprzyk PG, King CR;					
DR	WPI; 97-064831/06.					
DR	N-PSDB: T65007.					
PT	Single chain antibodies specific for erbB-2 protein, gp185 - with					
PT	labels or cytotoxin, useful for detection and treatment of tumour					
PT	cells expressing this protein					
PS	Claim 2; Columns 27-30; 28pp; English.					
CC	The present sequence represents a claimed single-chain antibody,					
CC	designated e21(Fv), which binds to erbB-2. Monoclonal antibody e21					
CC	was generated by immunising mice with N/erbB-2 cells overexpressing					
CC	the gp185 protein, removing spleen cells and producing hybridomas					
CC	by standard techniques. Messenger RNA coding for the anti-erbB-2					
CC	monoclonal antibody was isolated and converted to cDNA. Regions					
CC	coding for the heavy- and light- chain variable regions were then					
CC	amplified by PCR and joined via a sequence encoding a peptide					
CC	linker. The resulting single-chain antibody is useful for in vitro					
CC	diagnosis of tumour cells which overexpress the erbB-2 gp185					
CC	marker, e.g. breast, ovarian and non-small cell lung carcinomas,					
CC	and, when coupled to a cytotoxic agent, to treat such tumours.					
SC	Sequence 239 AA;					
Query Match	99.4%; Score 1264; DB 1; Length 239;					
Best Local Similarity	100.0%; Pred. No. 4.5e-82;					
Matches 239; Conservative	0; Mismatches 0; Indels 0; Gaps					
QY	1 MOLTSPAIMSASPEKRYTMTCSASSSVSNMHWYQOKSSTSKLWYDTSKLASGYPGRF	60				
DB	1 MOLTSPAIMSASPEKRYTMTCSASSSVSNMHWYQOKSSTSKLWYDTSKLASGYPGRF	60				
QY	61 SSGSGNSNSTLITSSWEADATYYCYOGSGYPFFGSGTKLEIGSTSGSKSSPGKGV	120				
DB	61 SSGSGNSNSTLITSSWEADATYYCYOGSGYPFFGSGTKLEIGSTSGSKSSPGKGV	120				
QY	121 QIQOGSVELVIRGALVYKLSCKASDENIKYIHWKQREQGLWIGWIHPENGNNVYDP	180				
DB	121 QIQOGSVELVIRGALVYKLSCKASDENIKYIHWKQREQGLWIGWIHPENGNNVYDP	180				
QY	181 KFGQKASTADTSSNNAATQLSSLTSEDTAVYYCASYYIYATYTWGGGTSTVSS	239				
DB	181 KFGQKASTADTSSNNAATQLSSLTSEDTAVYYCASYYIYATYTWGGGTSTVSS	239				

ID R45443 standard; Protein: 238 AA.  
AC R45443:  
DE 21-JUL-1994 (first entry)  
DE Sequence of the single chain anti-erbB2 antibody, Ab no.21.  
KW Single chain anti-erbB1 antibody; cancer therapy; prevention;  
KW monoclonal antibody.  
OS Synthetic.  
PN WO940136-A.  
PD 06-JAN-1994.  
PF 21-OCT-1992; U08545.  
PR 30-JUN-1992; US-906555.  
PA (MOLE-) MOLECULAR ONCOLOGY INC.  
PI Kaspitzky BG, King CR;  
DR WPI: 94-025878/03.  
DR N-PSDB: Q55181.  
PT Treatment of malignancies over-expressing ERB-2 - using at least  
PT 2 monoclonal antibodies which recognise different epitopes on  
PT gp185  
PS Example: Fig 8: 37pp. English.  
CC The source of human erbB-2 protein for the prodn. of antibodies no.  
CC 23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell  
CC engineered to express the human erbB-2 protein on its surface (N/  
CC erbB-2). Abs no. 21 and 23 are directed against the extracellular  
CC domain of gp185 erbB-2. Nude mice manipulated to produce rapidly  
CC growing tumours were used in a trial of the efficacy of the Abs. In  
CC animals given a combination of the 2 Abs, tumours completely  
CC regressed after 11 days.  
SQ Sequence 238 AA;

Query Match 99.1%; Score 1259; DB 1; Length 238;  
Best Local Similarity 100.0%; Pred. No. 1e-81;  
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QLTQSPALMSASPGKVTMTCSASSSVSNMHWYQOKSTSPKLMYDTSKLASGVPGRF 61  
DB 1 QLTQSPALMSASPGKVTMTCSASSSVSNMHWYQOKSTSPKLMYDTSKLASGVPGRF 60  
QY 62 GSGSGNSYSLTISSEMEDAATYYCYQSGYPFTFGSGTKLEIKSGTSGSGSSEKGTQ 121  
DB 61 GSGSGNSYSLTISSEMEDAATYYCYQSGYPFTFGSGTKLEIKSGTSGSGSSEKGTQ 120  
QY 122 LQOQSGVELYRGALVRLSKASDFNFKDYIHWVKORPEQGLEWIGWHPENGNTYVDK 181  
DB 121 LQOQSGVELYRGALVRLSKASDFNFKDYIHWVKORPEQGLEWIGWHPENGNTYVDK 180  
QY 182 FQGKASITADTSSNAAYLQLSLTSEDTAVYYCASYYYSAYYAMYWGQTSVTYSS 239  
DB 181 FQGKASITADTSSNAAYLQLSLTSEDTAVYYCASYYYSAYYAMYWGQTSVTYSS 238

RESULT 3  
R60780  
ID R60780 standard; Protein: 244 AA.  
AC R60780:  
DE 09-MAY-1995 (first entry)  
DE Fv(GP-2) immunosuppressive.  
KW MAb; monoclonal antibody; hybridoma; interleukin-2; IL-2;  
KW Fv; antibody variable region; GP-2; Fv(GP-2);  
KW immunosuppressive.  
OS Mus sp.  
PN EP-621338-A.  
PD 26-OCT-1994.  
PF 21-APR-1994; 106257.  
PR 21-APR-1993; JP-094491.  
PR 07-MAR-1994; JP-036065.  
PA (AJTN) AJINOMOTO KK.  
PI Hamura J, Kanayama Y, Nakazawa H, Shimamura T, Sugamura K;  
PI Takeshita T;  
DR WPI: 94-325948/41.  
DR N-PSDB: Q73678.  
PT Immunosuppressant polypeptide - has ability to block  
PT interleukin-2 response

PS Claim 8; Page 26-27; 37pp. English.  
CC MAb capable of binding to the gamma chain of the IL-2 receptor, and  
CC thus of blocking the IL-2 response, is produced by mouse hybridoma  
CC line GP-2 (FERM BP-4641). DNA encoding the variable region of  
CC this MAb was expressed in E. coli, yielding Fv(GP-2) with  
CC immunosuppressive activity.  
SQ Sequence 244 AA;

Query Match 65.3%; Score 829.5; DB 1; Length 244;  
Best Local Similarity 68.3%; Pred. No. 1.4e-51;  
Matches 164; Conservative 27; Mismatches 46; Indels 3; Gaps 2;

QY 3 LTPSPALMSASPGKVTMTCSASSSVSN-MHWYQOKSTSPKLMYDTSKLASGVPGRF 60  
DB 5 LTPSPALMSASPGKVTMTCSASSSVSNLHWYQOKRGSPPKLMYDTSKLASGVPGRF 64  
QY 61 GSGSGNSYSLTISSEMEDAATYYCYQSGYPFTFGSGTKLEIK-GSTSSGSGSSEKGTQ 119  
DB 65 GSGSGNSYSLTISSEMEDAATYYCHQHRSPFTFGAGTKLEIKVDKSGSGSSEKSTQ 124  
QY 120 VLOQSGVELYRGALVRLSKASDFNFKDYIHWVKORPEQGLEWIGWHPENGNTYVD 179  
DB 125 VKLESGSELYRPGASVRLSKASGTYTSTYMMHWKORHOGGLEWIGNITPGSGSTYVD 184  
QY 180 PFQKASITADTSSNAAYLQLSLTSEDTAVYYCASYYYSAYYAMYWGQTSVTYSS 239  
DB 185 EFKFSKATLTVDTSSSTAYMLSLTSEDSAVYYCTRSSRMWYVYAMDYWGQTSVTYSS 244

RESULT 4  
R12798  
ID R12798 standard; Protein: 241 AA.  
AC R12798:  
DE 16-SEP-1991 (first entry)  
DE OV33 light and heavy chains.  
KW Immunoglobulin; immunotoxin; antibody; variable region; PE40;  
KW pseudomonas exotoxin.  
OS Synthetic.  
FH Key  
FH Location/Qualifiers  
FT region 1..107  
FT /label= light chain  
FT /label= linker  
FT region 108..121  
FT /label= linker  
FT region 122  
FT /label= Heavy chain  
FT region 24..33  
FT /label= CDR1  
FT region 49..54  
FT /label= CDR2  
FT region 88..96  
FT /label= CDR3  
FT region 151..155  
FT /label= CDR1  
FT region 170..186  
FT /label= CDR2  
FT region 219..229  
FT /label= CDR3  
FT peptide 1..20  
FT /label= N-terminal sequencing  
FT peptide 121..134  
FT /label= N-terminal sequencing  
FT peptide US7463111-A.  
PN 28-MAY-1991.  
PD 11-JAN-1990; 133066.  
PR 11-JAN-1990; US-463111.  
PR (USSH) NAT INST OF HEALTH.  
PA Pastan I;  
PI Pastan I;  
DR WPI: 91-200877/27.  
DR N-PSDB: 012405.  
PT Rapid cloning of antibody genes as single chain immuno-toxins -  
PT by fusing genes with DNA encoding cytotoxic molecules, to kill  
PT target cells

PS Disclosure: Fig 3: 26pp: English.  
 CC The sequence comprises the L and H chains of murine Mab OYB3,  
 CC joined together via a linker. The sequence is encoded by clone  
 CC pOVB315-1 in which the 3' end of the DNA is linked to the 5' end  
 CC of the coding sequence for pE40, a Pseudomonas exotoxin lacking  
 CC domain I. The protein expressed by the clone can be used to kill  
 CC specific target cells.  
 SQ Sequence 241 AA;

Query Match 64.6%; Score 821.5; DB 1; Length 241;  
 Best Local Similarity 66.1%; Pred. No. 5, 1e-51;  
 Matches 158; Conservative 31; Mismatches 47; Indels 3; Gaps 3;

QY 3 LTGSPAIMSAPGEKVTMTCSASSVSNMHWYQOKSSTPKLWYDTSKLASGVPGRFG 62  
 DB 4 LTGSPAIMSAPGEKVTMTCSAISTVSNMHWYQOKSSTPKLWYDTSKLASGVPGRFG 63  
 QY 63 SSGGNSYSLSLTSSMEADAAATYYCQSGYPTFGSGTKLEIK-GSTSGSGKSGKGVQ 121  
 DB 64 SSGGNSYSLSLTSSMEADAAATYYCQSGYPTFGSGTKLEIKSGSGSGSKVDVQ 123  
 QY 122 LOOSGVELVKGALVTKLSCASDENIKDYIHWKORPEQGLEWIGWHPENGNTVYDPK 181  
 DB 124 LVESGGGLVDPGSGROISCAASGFTTSSFGMHWYQAPKPEGLWVAIT-SRGGNTIYAN 182  
 QY 182 -FOGKASITADTSSNNAIYQLSLTSEDTAVVYCASYYYSAYYAMWYQGTSTVYSS 239  
 DB 183 TVGKGFITSDNPKNTFLQMTSLRSDDTAMYCARSHYGYRYANDWYQGTITNGSS 241

RESULT 5  
 R94020 ID R94020 standard; Protein: 237 AA.  
 AC R94020:  
 DT 21-MAY-1996 (first entry)  
 DE Anti-erbB2 scFv.  
 KW Oncoprotein; erbB2; cell proliferation; tumour; cancer;  
 KW Intracellular antibody homologue; single chain antibody; scFv;  
 KW gene therapy.  
 OS Synthetic.  
 PN WO9607321-A1.  
 PD 14-MAR-1996.  
 PE 23-AUG-1995; 010740.  
 PR 06-SEP-1994; US-301339.  
 PR 06-JUN-1995; US-468252.  
 PA (UABR-) DAB RES FOUND.  
 PI Curjel DT, Deshane J;  
 DR N-PSDB: T17307/17.  
 DR N-PSDB: T17728.  
 PT Inhibition of proliferation or survival of, esp. malignant erbB2,  
 PT cells - by introducing nucleic acid mol. encoding antibody homologue  
 PT which is expressed and binds. pref. erbB2; protein intracellularly  
 PS Disclosure: Page 29-30; 48pp; English.  
 CC A nucleic acid comprises a first sequence encoding a signal peptide  
 CC (R94019) linked to a second sequence (T17728) encoding a single  
 CC chain Fv fragment (R94020) that binds a human erbB2 oncoprotein.  
 CC The anti-erbB2 scFv portion is obtained by PCR using e23scFv  
 CC plasmid as template. The signal peptide directs the scFv to the  
 CC endoplasmic reticulum. The nucleic acid is incorporated into a  
 CC plasmid or viral vector to facilitate expression of the scFv antibody  
 CC homologue within e.g. an epithelial carcinoma cell. Intracellular  
 CC expression of the homologue inhibits surface expression of erbB2 and  
 CC thereby inhibits cell proliferation and cell survival and decreases  
 CC tumorigenicity  
 SQ Sequence 237 AA;

Query Match 62.6%; Score 795.5; DB 1; Length 237;  
 Best Local Similarity 64.3%; Pred. No. 3, 3e-49;  
 Matches 155; Conservative 32; Mismatches 45; Indels 9; Gaps 2;

QY 1 MOLTOSPAIMSAPGEKVTMTCSASSVSNMHWYQOKSSTPKLWYDTSKLASGVPGRF 60

DB 3 LQLTSPAILIASPGEKVTMTCAIPSVSYMHYQOKPSSPEPWYTTSTNLASGVPAR 62  
 QY 61 SSGGNSYSLSLTSSMEADAAATYYCQSGYPTFGSGTKLEIKSTSGSGSKGKV 120  
 DB 63 SSGGNSYSLSLTSSMEADAAATYYCQSGYPTFGSGTKLEIKSTSGSGSKGKV 122  
 QY 121 QLOOSGVELVKGALVTKLSCASDENIKDYIHWKORPEQGLEWIGWHPENGNTVYDP 180  
 DB 123 QLOESGPEVVKPGSGMKISCKTSGYSTGTMTMHWYQSHGKNLEWIGLIPYNGDTINYD 182  
 QY 181 KFGKASITADTSSNNAIYQLSLTSEDTAVVYCA---STYYSAIYAMWYQGTSTVY 237  
 DB 183 KFGKATFTVDKSSSTAYMELLSTSEDSAVVYCARVDWYFDV-----WGAGTIVY 236  
 QY 238 S 238  
 DB 237 S 237

RESULT 6  
 W15185 ID W15185 standard; Protein: 237 AA.  
 AC W15185:  
 DT 05-JUN-1997 (first entry)  
 DE Single-chain anti-erbB2 antibody e23(Fv).  
 KW Single chain antibody; variable region; light chain; heavy chain;  
 KW breast cancer; ovarian cancer; non-small cell lung carcinoma;  
 KW immunodiagnostics; treatment; cytotoxic agent; erbB2.  
 OS Mus musculus.  
 OS Synthetic.  
 FH Key  
 FT region  
 FT 108..121  
 FT /label= linker  
 PN US567458-A.  
 PD 24-DEC-1996.  
 PE 07-OCT-1991; 772270.  
 PR 30-OCT-1991; US-772270.  
 PR 30-JUN-1992; US-906555.  
 PR 14-MAY-1993; US-061092.  
 PA (ARON-) ARONEX PHARM INC.  
 PI Bird RE, Kasprzyk PG, Kling CR,  
 DR WPI: 97-064831/06.  
 DR N-PSDB: T65006.  
 PT Single chain antibodies specific for erbB-2 protein, gp185 - with  
 PT labels or cytotoxin, useful for detection and treatment of tumour  
 PT cells expressing this protein  
 PS Claim 2: Columns 25-28; 28pp; English.  
 CC The present sequence represents a claimed single-chain antibody,  
 CC designated e23(Fv), which binds to erbB-2. Monoclonal antibody e23  
 CC was generated by immunising mice with N/erbB-2 cells overexpressing  
 CC the gp185 protein, removing spleen cells and producing hybridomas  
 CC by standard techniques. Messenger RNA coding for the anti-erbB-2  
 CC monoclonal antibody was isolated and converted to cDNA. Regions  
 CC coding for the heavy- and light- chain variable regions were then  
 CC amplified by PCR and joined via a sequence encoding a peptide  
 CC linker. The resulting single-chain antibody is useful for in vitro  
 CC diagnosis of tumour cells which overexpress the erbB-2 gp185  
 CC marker, e.g. breast, ovarian and non-small cell lung carcinomas,  
 CC and, when coupled to a cytotoxic agent, to treat such tumours.  
 SQ Sequence 237 AA;

Query Match 62.6%; Score 795.5; DB 1; Length 237;  
 Best Local Similarity 64.3%; Pred. No. 3, 3e-49;  
 Matches 155; Conservative 32; Mismatches 45; Indels 9; Gaps 2;

QY 1 MOLTOSPAIMSAPGEKVTMTCSASSVSNMHWYQOKSSTPKLWYDTSKLASGVPGRF 60  
 DB 3 LQLTSPAILIASPGEKVTMTCAIPSVSYMHYQOKPSSPEPWYTTSTNLASGVPAR 62  
 QY 61 SSGGNSYSLSLTSSMEADAAATYYCQSGYPTFGSGTKLEIKSTSGSGSKGKV 120

```

Db      63  SGGSGSTSYSLTVSRVDAEDAAATYYCOQMSRSPPTFEGGSKLEIKGSTSGSGKSGSEKGV 122
QY      121  QLOOSGEVLVKGALVYKLSKASDENIKDYIHWVKORPEOGLEWIGWIPENGNTVYDP 180
      123  QLOESGEVYKPGGSMKISCKTSGTGTHTMMVKOSHKNLEWIGLIPYNGDINQV 182
QY      181  KFGGKASTIDTSSNAAYLQSLTSEDVAVYYCA---STYYSAAYAAATYWGOGTSVTV 237
      183  KFGKATFTVDKSSSTAYMELSLTSEDSALYYCARRVTDVYDVFY-----WGAGTIVTV 236
QY      238-S 238
      237 S 237

RESULT  7
R45442
ID      R45442 standard; Protein: 236 AA.
AC      R45442;
DT      21-JUL-1994 (first entry)
DE      Sequence of the single chain anti-erbB2 antibody, Ab no.23.
KW      Single chain anti-erbB1 antibody; cancer therapy; prevention;
      KM monoclonal antibody.
OS      Synthetic.
PN      W09400136-A.
PD      06-JAN-1994.
PF      21-OCT-1992; U08545.
PR      30-JUN-1992; US-906555.
PA      (MOLE-) MOLECULAR ONCOLOGY INC.
PI      KASPRZYK PC, KING CR.
DR      WPI; 94-025878/03.
N-PSDB; Q55180.
PT      Treatment of malignancies over-expressing ERB-2 - using at least
PT      2 monoclonal antibodies which recognise different epitopes on
PT      gp185
PS      Example; Fig 7, 37pp; English.
CC      The source of human erbB-2 protein for the prodn. of antibodies no.
CC      23 and 21 (Q55180/R45442; Q55181/R45443) is a NIH/3T3 cell
CC      engineered to express the human erbB-2 protein on its surface (N/
CC      erbB-2). Abs no. 21 and 23 are directed against the extracellular
CC      domain of gp185 erbB-2. Nude mice manipulated to produce rapidly
CC      growing tumours were used in a trial of the efficacy of the Abs. In
CC      animals given a combination of the 2 Abs, tumours completely
CC      regressed after 11 days.
SQ      Sequence 236 AA;

```

```

Query Match      62.2%; Score 790.5; DB 1; Length 236;
Best Local Similarity 63.9%; Pred. No. 7.5e-49;
Matches 154; Conservative 32; Mismatches 46; Indels 9; Gaps 2;

```

```

QY      1  MOLTQSPAIMSASPEKVTMTCSASSSVSMHWYQOKSSTSPKLMVYDTSKLASGVPGRF 60
      2  LQLTQSPAILISASPEGEVYMTCRATPVSSTMHYQOKPGSSPKFWITTNLSGVPARF 61
QY      61  SGGSGSGSYSLTSSMEADAAATYYCYQSGGYPTFEGSGTKLEIKGSTSGSGKSGSEKGV 120
      62  SGGSGSGSYSLTVSRVDAEDAAATYYCOQMSRSPPTFEGGSKLEIKGSTSGSGKSGSEKGV 121
QY      121  QLOOSGEVLVKGALVYKLSKASDENIKDYIHWVKORPEOGLEWIGWIPENGNTVYDP 180
      122  QLODSGEVYKPGGSMKISCKTSGTGTHTMMVKOSHKNLEWIGLIPYNGDINQV 181
QY      181  KFGGKASTIDTSSNAAYLQSLTSEDVAVYYCA---STYYSAAYAAATYWGOGTSVTV 237
      182  KFGKATFTVDKSSSTAYMELSLTSEDSALYYCARRVTDVYDVFY-----WGAGTIVTV 235
QY      238 S 238
      236 S 236

RESULT  8

```

```

R15443
ID      R15443 standard; Protein: 240 AA.
AC      R15443;
DT      25-FEB-1992 (first entry)
DE      Single chain Fv from pSCV1 for blocking HRV binding to ICAM-1.
KW      antigen-binding fragment; inflammation; auto-immune disease.
OS      Homo sapiens.
PN      EP-459577-A.
PD      04-DEC-1991.
PF      25-MAY-1991; 201243.
PR      01-JUN-1990; US-532001.
PA      (MERI) MERCK & CO INC.
PI      Colono RJ, Condra JH, Tomassini JE, Sardana VV;
DR      WPI; 91-355850/49.
PT      Microbially expressed portions of monoclonal antibody - can block
PT      attachment of rhinovirus ligands to inter-cellular adhesion
PT      molecule (ICAM-1)
PS      Claim 5; Page 21; 28pp; English.
CC      The Fv fragment encoded by pSCV1 was able to protect cells against
CC      infection by HRV-14. It was less active than modified Fv fragments
CC      which included a Gly-Gly-Gly-Ser flexible linker between the L
CC      and H chains as either a monomer or a dimer.
CC      See also R15437-R15442.
SQ      Sequence 240 AA;

```

```

Query Match      60.0%; Score 763; DB 1; Length 240;
Best Local Similarity 61.6%; Pred. No. 6.5e-47;
Matches 149; Conservative 33; Mismatches 50; Indels 10; Gaps 3;

```

```

QY      3  LTQSPAIMSASPEKVTMTCSASSSVS-NMHWYQOKSSTSPKLMVYDTSKLASGVPGRF 61
      4  LQSPATLSVTPEGDSVSLSCRASQGISNLMHWYQOKSHSPRLILKHAQSISGIPSRFS 63
Db      4  LQSPATLSVTPEGDSVSLSCRASQGISNLMHWYQOKSHSPRLILKHAQSISGIPSRFS 63
QY      62  GSGSGSYSLTSSMEADAAATYYCYQSGGYPTFEGSGTKLEIKGSTSGSGS-----SFG 117
      64  GSGSGDFTLSINSVTEDEFGMFCQOSNSWPTFEGGTKLEIKRADAPTVSIEPPSSE 123
Db      64  GSGSGDFTLSINSVTEDEFGMFCQOSNSWPTFEGGTKLEIKRADAPTVSIEPPSSE 123
QY      118  KGVQLQOSGEVLVKGALVYKLSKASDENIKDYIHWVKORPEOGLEWIGWIPENGNTV 177
      124  EHVQLQOSGAEIVKPKASVYKLSCTAGFINKDTYIHMVKORPEOGLEWIGRIDPANDNTI 183
QY      178  YDPKFGKASTIDTSSNAAYLQSLTSEDVAVYYCASYYSAAYAAATYWGOGTSVTV 237
      184  YDPKVGKATMTADTSSNAAYLQSLTSEDVAVYYCTTSGYFA-----YMGQGLTVTV 238
Db      184  YDPKVGKATMTADTSSNAAYLQSLTSEDVAVYYCTTSGYFA-----YMGQGLTVTV 238
QY      238 SS 239
      239 SA 240

```

```

RESULT  9
W25784
ID      W25784 standard; Protein: 240 AA.
AC      W25784;
DT      22-DEC-1997 (first entry)
DE      Anti-gp130 antibody derived scFv.
KW      Single chain Fv; scFv; antibody; E. coli; anti-T3 antibody;
      KM anti-gp130 antibody; inclusion body; chaperonin.
OS      Synthetic.
PN      J09220092-A.
PD      26-AUG-1997.
PF      15-FEB-1996; 027622.
PR      15-FEB-1996; JP-027622.
PA      (TOYT) TOSOH CORP.
DR      WPI; 97-474306/44.
N-PSDB; T91615.
PT      Producing single chain Fv antibody in Escherichia coli - by
PT      expression in an inclusion body, followed by protein folding or by
PT      co-expression with a chaperonin as a soluble fraction
PS      Claim 4; Page 6-7; 9pp; Japanese.
CC      The sequences given in W25783-84 represent single chain Fv (scFv)
CC      antibodies which are produced in E.coli. The scFv's are derived

```

FT	region	5
FT		/

PT specific killing comprising an anti

anti-Fc receptor antibody

PR a target cell  
PS Example 8; Fig 40; 57pp; English.  
CC This sequence represents the construct 321, which is a fusion between an  
CC anti-CEA antibody and a multispecific single chain antibody designated  
CC H22. The antibody can be used in the method of the invention for inducing  
CC antibody-dependent cellular cytotoxicity (ADCC) against a tumour cell  
CC which is characterised by overexpression of HER 2/neu or epidermal growth  
CC factor receptor (EGFR), comprises contacting the tumour cell with a  
CC multispecific protein molecule (preferably a single chain antibody)  
CC comprising: (a) an anti-FC receptor antibody or an antigen binding  
CC fragment; (b) a portion which binds to HER 2/neu; and (c) a portion which  
CC binds to EGFR. The method can be used for treating cancers especially  
CC breast cancer or ovarian cancer. The multispecific antibody can also  
CC be administered prophylactically to vaccinate a subject against infection  
CC by a target cell.  
SQ Sequence 553 AA;

Query Match 59.2%; Score 752; DB 1; Length 553;  
Best Local Similarity 60.9%; Pred. No. 9.4e-46;  
Matches 154; Conservative 27; Mismatches 46; Indels 26; Gaps 7;

QY 1 MOLTOSPALMSAPGKVMTCSSASSV--SN---MHMYQOKSSTSPKLMWYDTSKLA 53  
DB 156 IQLTOSPSLSASVGDVRIITCKSSOSVLYSNQKNYLMYQOKPEKAKLLIYMASTRE 215  
QY 54 SCVYGRFSGSGSGSNYSLTISSEMEADATYYCYQGSY--PFTFGSGTKLEIKGSTGS 111  
DB 216 SGVPSRFSFGSGSDPTFTFTSLQPEDATYYCHQ---YLSMTPEQGRKVEIKSSCSGG 272  
QY 112 GRSSEKGYQLQOSGYELVYGALVKLSKASDENIKDYIHWVKORPEGLWIGTWIHP 171  
DB 273 GGGSD---IKLQOSGALVRSYVLSCTASGFFNIKDSYMHMLRQGPQGLEWIGWIDP 329  
QY 172 ENGNTYVDKFGKASTITADTSSNAAYLQLSLTSEDTAVYVC-----ASYYYSAAYAM 226  
DB 330 ENGDTYARKFGKATFTTDTSSNAYLQLSLTSEDTAVYVCNEGTPTGPYTFD----- 385  
QY 227 YWGGGTSTVTS 239  
DB 385 -YWGQGTITVTS 396

RESULT 12  
ID W02191  
AC W02191:  
DT 13-NOV-1996 (first entry)  
DE 18-2-3/TRI202 single chain binding protein.  
KW Antibody engineering; monoclonal antibody; Mab;  
KW single chain binding protein.  
OS Chimeric Mus musculus;  
OS Chimeric synthetic.  
PN US518889-A.  
PD 21-MAY-1996.  
PE 02-SEP-1986; 902971.  
PR 02-SEP-1986; US-902971.  
PR 02-SEP-1987; US-092110.  
PR 15-JAN-1988; US-299617.  
PR 25-APR-1990; US-512910.  
PR 01-APR-1993; US-040440.  
PR 06-JUN-1995; US-468988.  
PA (ENZO-) ENZON LABS INC.  
PI Bird RE, Hardman K, Ladner RC;  
DR WPI: 96-259060/26.  
DR N-PSDB: T36463.  
PT Immunassay using single chain antigen binding mol. - as replacement  
PT to engineer, more stable and less expensive  
PS Example 9; Fig 39A-B; 78pp; English.  
CC Computer-designed construct 18-2-3/TRI202' (W02191), encoded by  
CC the DNA sequence given in T36463, comprises the variable regions

CC of the light and heavy chains of anti-fluorescein monoclonal  
CC antibody (Mab) 18-2-3 linked by a peptide designed to fit into a  
CC groove on the backside of the variable domain structure. The DNA  
CC construct was inserted into vector PGX3703 and introduced into E.  
CC coli. 18-2-3/TRI202' was produced as a single chain molecule. It  
CC exhibited biological binding activity equivalent in specificity  
CC and affinity to that of the original Mab.  
SQ Sequence 239 AA;

Query Match 59.1%; Score 751.5; DB 1; Length 239;  
Best Local Similarity 62.2%; Pred. No. 4.1e-46;  
Matches 148; Conservative 30; Mismatches 55; Indels 5; Gaps 3;

QY 3 LTQSPALMSAPGKVMTCSSASSVSN--MHMYQOKSSTSPKLMWYDTSKLAGVPGRF 60  
DB 5 LTQSPALMSAPGKVMTCSSASSVSSYLHWYQOKSGAPKLMWYDTSNLAGSVPARF 64  
QY 61 SGGSGNSYSLTISSEMEADATYYCYQGSYPFTFGSGTKLEIKGSTSGSKSGKGV 120  
DB 65 SGGSGTSTSLTISSEMEADATYYCQYSGYPLTFGAGTKLEIKSGSKSG--SEKST 122  
QY 121 QLQOSGYELVYGALVKLSKASDENIKDYIHWVKORPEGLWIGTWIHPENGWYDTP 180  
DB 123 QLKESGYPVLVAPDSGLSTCTVSGFSLTNYGVHWYQPPGKLEWIGVIM--AGGNTNNS 181  
QY 181 KFGKASTITADTSSNAAYLQLSLTSEDTAVYVCASYYYSAAYMYWGGTSTVTS 238  
DB 182 ALMSLTSLSKNSKQVFLKKNLSLNDPTATYYCAKRGRIFFYAMDMWGGTSTVTS 239

RESULT 13  
ID R43679  
AC R43679:  
DT 23-MAY-1994 (first entry)  
DE Single chain polypeptide with affinity for fluorescein.  
KW Monoclonal antibody; Mab; affinity; binding; antigen; diagnostics;  
KW therapy; imaging; purification; biosensors.  
OS Synthetic.  
PN US5260203-A.  
PD 09-NOV-1993.  
PE 02-SEP-1986; 902971.  
PR 02-SEP-1986; US-902971.  
PR 02-SEP-1987; US-092110.  
PR 19-JAN-1989; US-299617.  
PR 25-APR-1990; US-512910.  
PA (ENZO-) ENZON INC.  
PI Bird RE, Hardman K, Ladner RC;  
DR WPI: 93-367875/46.  
DR N-PSDB: O51540.  
PT Single chain polypeptide for binding antigen - comprising light  
PT and heavy chain antigen binding portions linked by peptide linker  
PS Example 9; Figure 39; 78pp; English.  
PR The single chain polypeptide is derived from the mature light and  
CC heavy chains of a monoclonal antibody (Mab) and has affinity  
CC for a given antigen (Fluorescein). It comprises a first  
CC polypeptide comprising the antigen binding portion of of the light  
CC chain variable region of an antibody and a second polypeptide  
CC comprising the antigen binding portion of the heavy chain variable  
CC region of an antibody and at least one peptide linker linking the  
CC first and second polypeptide chains. The resulting single chain  
CC polypeptide can be used in diagnostics, therapy  
CC (in vivo and in vitro), imaging, purifications and biosensors.  
CC This particular single chain binding molecule was designated  
CC 18-2-3/TRI202' and contains one linker peptide.  
SQ Sequence 239 AA;

Query Match 58.9%; Score 748.5; DB 1; Length 239;  
Best Local Similarity 62.2%; Pred. No. 6.7e-46;  
Matches 148; Conservative 29; Mismatches 56; Indels 5; Gaps 3;





